Vision Science Symposium

A Tribute to Gordon G. Heath

Indiana University
1988

Digitized by the Internet Archive
in 2015

Vision Science Symposium

A Tribute to Gordon G. Heath

Indiana University
1988

The symposium and this volume are dedicated to
Gordon G. Heath as a tribute to his inspiring and
enduring contributions to Physiological Optics at
Indiana University

Bell Optical laboratories congratulates Dr. Gordon Heath on his thirty-three years service to Indiana University and to Optometry. We are happy to participate in this tribute.

—Mr. J. Hank Zobrist
Vice President

CIBA Vision Corporation is pleased to join with the many friends and admirers of Dr. Gordon Heath in honoring him for his many outstanding years in Optometric Education.

—Dr. B. J. Shannon
Executive Director
Professional Services
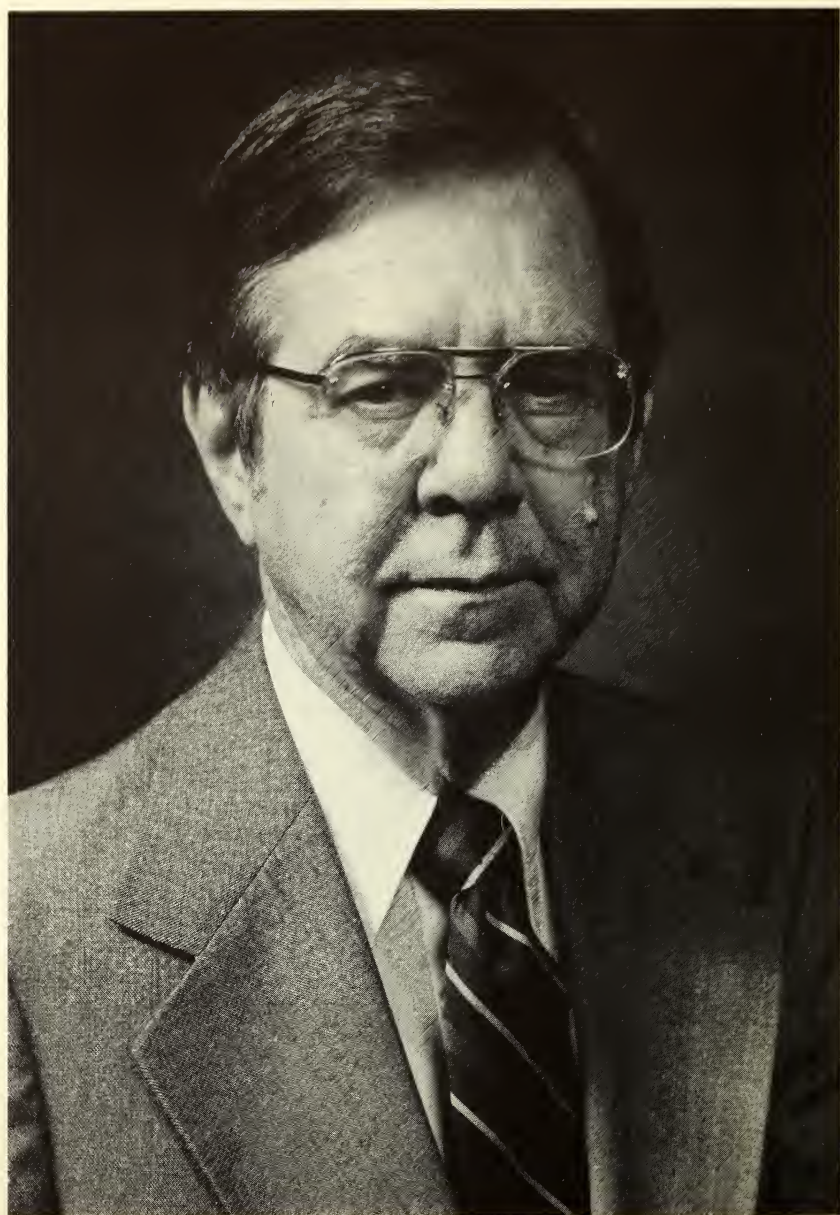
Gordon Glenn Heath was born in Sultan, Washington, September 22, 1922, but always considered himself a native Californian because he and his family moved there when he was only 5 years old. He enrolled in the University of California at Berkeley in 1940 to pursue a degree in chemical engineering, but the outbreak of World War II altered that plan. He joined the U.S. Naval Reserve and, in its V-12 program, attended Los Angeles City College, Peru (Nebraska) State Teachers College, and Northwestern University in Chicago, where he graduated from Midshipmen's School with a commission as Ensign. He was immediately sent to New Guinea where he joined his ship, the U.S.S. Carter Hall (LSD-3), just in time to participate in the invasion of Leyte Gulf in the Philippines. Later, after other invasions and voyages to many Pacific ports, the Carter Hall was sent soon after war's end to Shanghai, where Heath, still an Ensign but by now senior to the ship's other 25 officers, was appointed its Commanding Officer, a post he held until returning to the U.S. for discharge from active duty. Ironically, it was while on terminal leave before rejoining civilian life that he finally was promoted to Lt. (j.g.).

In 1946 Professor Heath resumed his education, entering the University of Southern California. He was influenced to study optometry by a boyhood friend whom he joined at Los Angeles College of Optometry in 1948, where he earned the B.V.S. degree in 1950 and the O.D. degree in 1951. He entered the graduate program in physiological optics at the University of California, Berkeley in the fall of that year, later receiving the M.S. and Ph.D. degrees.

Professor Heath has been at Indiana University since 1955. He served as director of the optometry clinic from 1955 to 1960 and director of the graduate program in physiological optics from 1960 to 1970, at which time he became Director of the Division of Optometry. Under his leadership, the Division was elevated to School status in 1975 and he was appointed Dean, a position he held until July 31, 1988, when he resumed his faculty post.

Dr. Heath served as president of the Association of Schools and Colleges of Optometry from 1963 to 1965, and president of the American Academy of Optometry for the years 1983 and 1984. He has held appointments as research consultant to the Office of the Surgeon General of the Army; member of the National Academy of Sciences/National Research Council Committee on Vision; the Advisory Council for Health Professions, HEW; and the National Advisory Eye Council, National Eye Institute, NIH.

Dr. Heath's research interests have centered on color vision and photoreceptor directionality. He has authored numerous papers in these and other research and professional subjects. He is highly respected generally for both his scientific and professional contributions; but those who have been his graduate students have for him a very special affection. The tribute of Professor Tony Adams, one of his graduate students, exemplifies the high degree of esteem in which he is held by all: "Gordon Heath has been the mentor of many, both formally and informally. I am most grateful to be among his many students and friends. Gordon's encouragement, guidance, patience and insight in vision research have placed him among the 'most admired' by students of visual science. He stands very tall without casting a threatening shadow. Of course a symposium dedicated to him is most appropriate and I am pleased to join in recognizing his exceptional contributions over the years."

TABLE OF CONTENTS

Symposium Participants

Back Row (l to r): J. A. Worthey, M. M. Uslan, J. K. Hovis, J. R. Zahn, D. L. McCulloch, A. J. Adams, I. L. Bailey, D. G. Pitts, D. A. Goss, G. L. Trick

Front Row (l to r): J. E. Bailey, R. W. Massof, G. G. Heath, R. F. Applegate

CONTRIBUTORS

Anthony J. Adams, O.D., Ph.D., University of California at Berkeley.

Raymond F. Applegate, O.D., Ph.D., University of Texas Health Science Center.

Ian L. Bailey, O.D., M.S., University of California at Berkeley.

James E. Bailey, O.D., Ph.D., Southern California College of Optometry.

David A. Goss, O.D., Ph.D., Northeastern State University College of Optometry.

Jeffery K. Hovis, O.D., Ph.D., University of Waterloo.

Robert W. Massof, Ph.D., Johns Hopkins Medical Institutions.

Daphne L. McCulloch, O.D., Ph.D., University of Toronto.

Donald G. Pitts, O.D., Ph.D., University of Houston College of Optometry.

Gary L. Trick, Ph.D., Washington University.

Mark M. Uslan, M.Ed., M.S., American Foundation for the Blind.

T. David Williams, O.D., Ph.D., University of Waterloo.

James A. Worthey, Ph.D., National Institute of Standards and Technology.

Joseph R. Zahn, Ph.D., Director of Low Vision Services, Baptist Eye Institute, Knoxville, Tennessee.

PREFACE

The symposium from which this volume is drawn was held on September 9, 1988. This was one of a series of events that were organized to pay tribute to Professor Gordon G. Heath on the occasion of his retirement from the deanship at the School of Optometry, Indiana University. The idea of bringing together former students of the graduate program in physiological optics for this purpose was the brainchild of Ray Applegate, Robert Massof, and Gary Trick. They were also largely responsible for arranging the symposium program. The willing response of those asked to contribute papers and the number of graduates of the physiological optics program who attended are true testimonials to the great respect Dr. Heath has earned as mentor, and the high regard he has earned as friend.

The fourteen papers presented in this symposium volume cover a wide range of subjects—so wide that they resist grouping on a topical basis. Their common factor is simply that their authors are all graduates of the program in physiological optics that was chaired by Dr. Heath throughout the decade of the 60's prior to his becoming dean.

The first six papers represent basic research with specific emphasis on vision and the eye. The J. Bailey and the Applegate-Massof papers concern retinal directional sensitivity, a topic of special interest to Professor Heath who devoted a sabbatical leave to research on this topic at the Institute for Perception in Soesterberg, Holland in 1969-70. The next four papers, by Adams, Trick, Hovis, and Massof deal with research in color vision, a longstanding interest of Dr. Heath stemming from his experiences as a graduate student under Gordon L. Walls at Berkeley. Each of the six papers contains a review of the research literature in its specific topical area.

The next six authors deal with various ways in which research may be applied. Worthey discusses, with wit as well as insight, the optics of veiling reflections. Pitts comprehensively considers the effects of environmental UV radiation on the eye. Goss postulates a causative mechanism for juvenile myopia based on animal experiments concerning refractive effects of retinal image degradation. McCulloch reports on clinical evaluation of vision in aphakic children as an application of visually evoked potentials. Two authors deal with improvements in vision testing procedures: Williams applies computer techniques to the quantification of visual field measures; I. Bailey presents the problems and difficulties of measuring and reporting visual acuities, and recommends a simplified method of standardizing them.

The final two authors present practical application of remedial work with the visually impaired: Zahn describing the design and operation of a model vision rehabilitation service center; Uslan discussing history and current technological use of devices for orientation and mobility of the blind or severely visually impaired.

In addition to those who originated the idea of a symposium by and for those who came under Heath's tutelage as graduate students, the contributions of others who participated should be acknowledged: Gloria Cochran, Dan Gerstman, Sue Gilmore and Ron Jensen, all members of the optometry administrative staff, made certain that preparations and arrangements for all events went smoothly. Those members of the I.U. optometry faculty served as moderators at the symposium sessions: Arthur Bradley, Robert Devoe, S. Lee Guth, and Larry Thibos. We are also indebted to Bell Optical Laboratories and to Ciba Vision Corporation for helping underwrite publication of the symposium papers.

A wealth of memories of graduate student days evolved among those who gathered for this event. The breadth of contributions by those who presented papers evokes assurance that those who passed that way during the Heath years will continue productively to expand his influence on visual science and vision care for a long time to come.

William R. Baldwin
July 29, 1989

FACTORS DETERMINING HUMAN RECEPTOR ORIENTATION
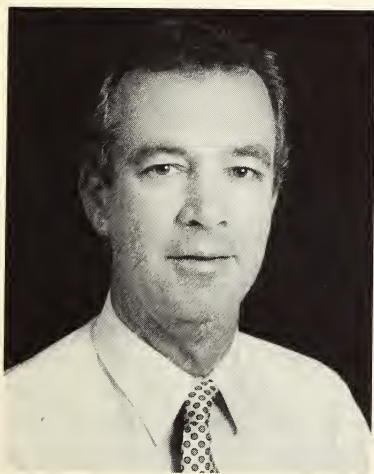
James E. Bailey
Associate Professor of Visual Science and Optometry at Southern California College of Optometry, Fullerton, California where he is Chair of the Department of Basic and Visual Sciences.

ABSTRACT. Evidence is emerging from psychophysical measurements of the Stiles-Crawford effect (SCE) that a combination of forces (phototropic, center of retinal sphere orienting, traction, packing) influence photoreceptor alignment. This evidence is reviewed, and new SCE data is reported demonstrating anomalous photoreceptor orientation in a variety of colobomatous eyes.

FACTORS DETERMINING HUMAN PHOTORECEPTOR ORIENTATION

I am grateful for the invitation to participate in this visual science symposium honoring Gordon Heath. It was he who got me interested in the problem of the directional sensitivity of the retina, generally accepted as the explanation of the Stiles-Crawford (S-C) effect. During a sabbatical leave in the Netherlands he worked with Pieter Walraven on psychophysical studies of photoreceptor orientation and directional sensitivity.[1,2] Following this, he encouraged me to pursue an investigation of photoreceptor directional sensitivity for my doctoral research.[3,4] Afterwards, he helped get me to Walraven's laboratory where I continued studies of the Stiles-Crawford effect in collaboration with J.J. Vos.[5] I shall be forever grateful to Gordon Heath for these opportunities.

My contribution to this symposium is to review evidence indicating that photoreceptor orientation is influenced by a combination of several forces or factors. This evidence has emerged from psychophysical studies of the S-C effect of normal observers and those with ocular anomalies. Additionally, I will present some results of investigations on this subject, recently conducted in collaboration with Drs. Jay Enoch and Vasudevan Lakshminarayanan during a sabbatical leave I spent in their laboratory at the School of Optometry, University of California, Berkeley.

The Stiles-Crawford Effect

The sensitivity of the eye to light depends upon the point in the pupil through which the light enters the eye. Sensitivity is generally greatest near the center of the pupil but progressively decreases for more eccentric entry points. This effect was discovered by Stiles and Crawford[6] and is most pronounced when measured under conditions favoring cone-mediated photopic vision.[7] A S-C function is depicted in the top panel of Fig. 1, illustrating the falloff in sensitivity from the peak of sensitivity near the center of the pupil. The lower panels of Fig. 1 depict current ideas of the factors contributing to the S-C effect. These are reviewed by Enoch and Tobey[8] and will be summarized in the next section.
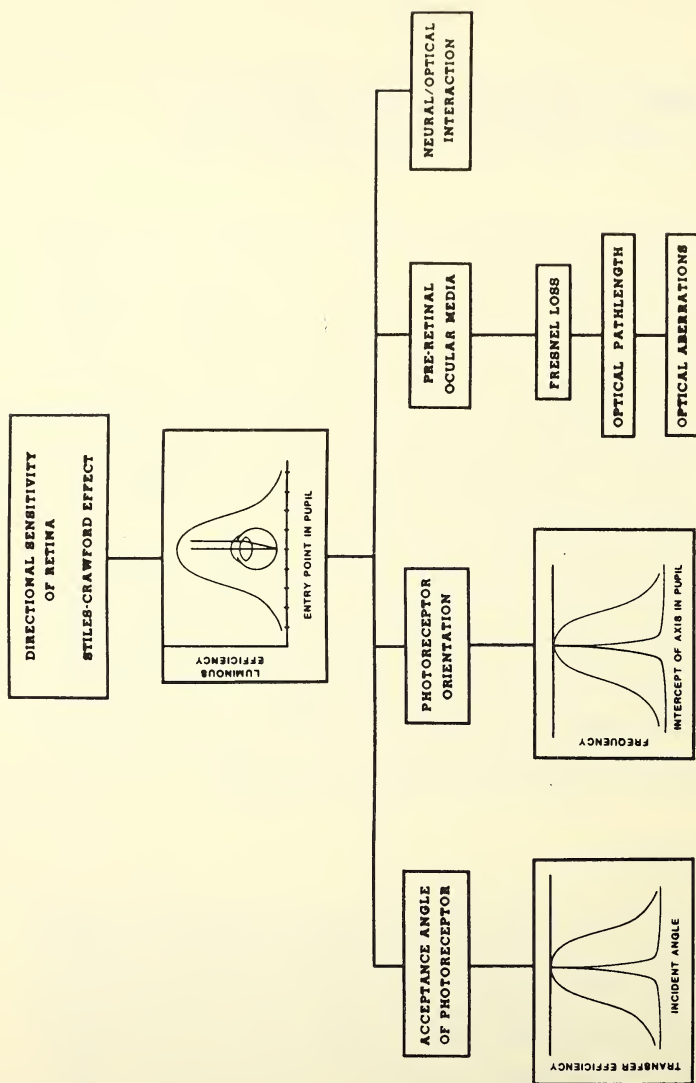
Fig 1. An illustration of the photopic Stiles-Crawford effect and factors affecting the variation of sensitivity as the point of entry of light in the pupil is varied. The concept of the acceptance angle of a photoreceptor (lower left panel) is used to describe its directional sensitivity, the efficiency with which energy entering the receptor at a given angle of incidence is transferred to the site of its absorption by photopigments. Curve width represents acceptance angle of receptor. Relative to flat line at top (no directional sensitivity) directional sensitivity increases as acceptance angle becomes progressively narrower. Photoreceptor orientation (bottom panel, second from left) is described by a distribution (assumed normal) of the frequency that the orientation axes of photoreceptors intercept a given location in the pupil. The flat line at top represents no preferred orientation (random) and progressively narrower distributions indicate greater precision (less dispersion) of alignment. See text for additional details.

Different points of pupil entry result in different angles of incidence of light on the retina. Variation of sensitivity with angle of incidence is attributed to an inherent directional sensitivity of the retina. As depicted in the lower left panel of Fig. 1, photoreceptors are directionally sensitive. However, the extent to which the directionality of the retina is indicative of individual photoreceptor directional sensitivity is not yet clear.[3,9,10,11] If individual cones have acceptance angles narrower than the S-C function (narrow curve in lower left panel), then the S-C effect can be considered the result of some variation in the intercepts of the receptors' orientation axes. This is represented by the broad curve in the panel second from the left at the bottom that matches the S-C function in the top panel. If, on the other hand, the acceptance angle of receptors is as broad as the overall S-C function (broad curve, lower left panel), then there would be little variation in where the orientation axes of receptors intercept the pupil (narrow curve, second from left at bottom). If both the acceptance angle and the distribution of receptor orientations are broad, the S-C effect function would be much broader, approximating a flat line.

Pre-retinal media factors, affecting light transmission and imagery in the eye, could also influence the S-C effect. To some extent this depends on the chosen stimulus characteristics. These media factors are listed in the lower panels, third from the left in Fig.1 and are discussed by Weale.[12] Finally, as listed in the lower right panel of Fig. 1, neural and optical interactions between elements in the retinal area sampled could also modify the group response.

Photoreceptor Alignment Mechanisms

Whatever relationships exist between photoreceptor directionality and orientation, the peak of the S-C function can be considered to represent the general orientation tendency of the group of receptors sampled by the measurement. The width of the function reflects variation in directional properties and dispersion in alignment of the sampled elements. Alignment of directionally sensitive receptors towards the pupil ensures maximum efficiency in utilization of light entering the eye. Measurements of the S-C function at different retinal locations offers a means to characterize mechanisms defining receptor orientation. The peak of the S-C function measured at fixation is usually located near the center of the pupil,[13,14] remains centered when measurements are taken over a wide retinal area,[15,16] and is remarkably stable over time.[17] Numerous reports have described recovery of the S-C function following an initial disturbance (flattening or location within pupil displaced, or both) due to pathology.[8] These findings imply the existence of an active mechanism, perhaps phototropic,[18] for maintaining alignment of photoreceptors with the center of the pupil. Ad-

ditional studies, which I will now review, support the notion of a photo-
tropic orienting force, but also indicate that other forces act to influence
receptor alignment.

The location of the peak of the S-C effect has been found to be shifted
in the direction of a displacement of the pupil due to injury.[19] A schematic
of the anomalous pupil and a measurement of the S-C effect in this eye
are shown in the upper half of Fig.2. The S-C function on the right, obtained
from an oblique traverse of the dilated eccentric pupil, is displaced in the
direction of the displaced pupil. Maximum efficiency was found to be near
the location of the natural eccentric pupil. In a subsequent investigation,
the subject (Bonds) was fitted with a contact lens having a centered artificial
pupil. This study was undertaken to determine if there was an active
alignment mechanism responsible for orienting photoreceptors.[20] The S-C
function obtained after the lens was worn for several days with the pupil
dilated is shown in the lower half of Fig.2. Clearly, there is a temporalward
shift in the peak location of the S-C effect, away from the natural eccentric
pupil and toward the centered artificial pupil, indicating an active, pho-
totropic pupil-orienting mechanism.

Similar investigations were conducted by Enoch and Birch.[21] Enoch
wore a contact lens with a displaced aperture in an attempt to produce a
comparable amount of translation of the peak of the S-C effect. The results
are summarized in Fig. 3A-C. Fig. 3A shows the location in the pupil of
the peaks of the S-C effect at different retinal locations. For points on
temporal retina, corresponding to nasal visual field, the peaks of the S-C
functions clustered near the center of the pupil; nasal retina points inter-
cepted the pupil in a more nasalward direction. Receptors across the tested
area of retina are affected by more than one orienting force.

To test for the presence of an active pupil-orienting mechanism, a
painted iris contact lens with a 2mm aperture was fitted. The aperture was
located 2mm temporalward in the pupil. Results of S-C effect measure-
ments after the lens was worn for 12 days are shown in Fig. 3B. Only the
peak locations of points on the temporal retina were displaced, but opposite
to the direction of the displaced artificial pupil. Increasing the diameter of
the aperture to 3mm and displacing the aperture 2.5mm temporal to the
pupil center produced the results shown in Fig. 3C. Points on the temporal
retina were now oriented toward the displaced aperture while nasal ori-
enting elements remained unaffected. Multiple factors are influencing re-
ceptor orientation in this eye. There is an indication of a mechanical (traction)
force with an origin near the optic nerve and an active light-induced force.
The 2 mm aperture presumably decreased the effect of the active pupil-
orienting force, causing the mechanical force to dominate (Fig. 3B). In-
creased light induced a graded alteration of alignment across the retina
toward the displaced aperture (Fig. 3C). The S-C effects measured for the
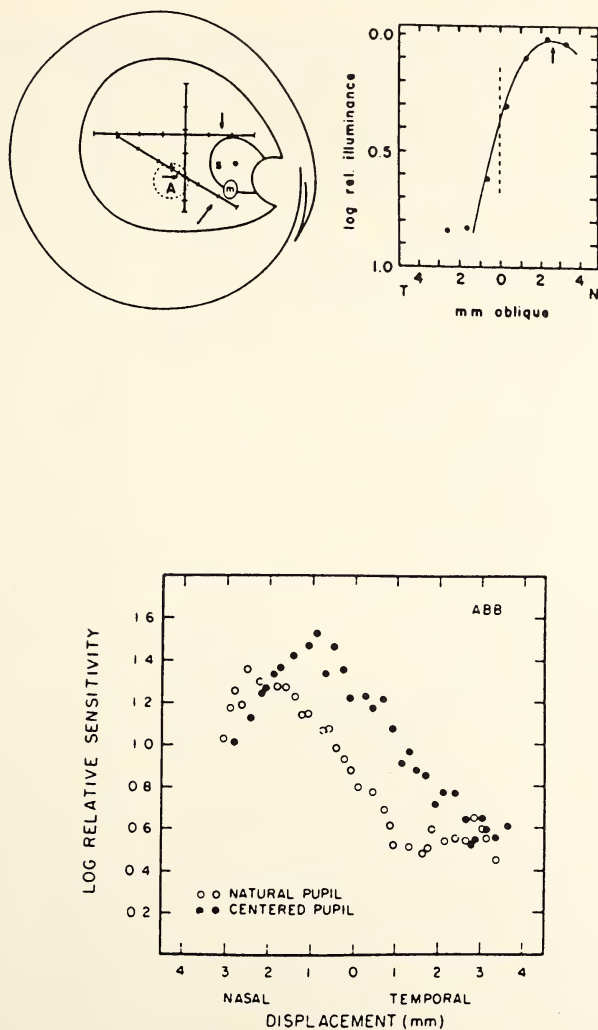
Fig 2. Stiles-Crawford effect data obtained from the amblyopic eye of A.B. Bonds (from Bonds and MacLeod, 1978; Applegate and Bonds, 1981). The upper panels show a diagram of the natural (small oval with dot at center) and dilated eccentric pupil (surrounding natural pupil) in relation to limbus (outermost contour), and the Stiles-Crawford function resulting from a traverse along oblique line. Eccentric pupil, adjacent notch, and nearby disturbance of medial limbus were result of trauma. Symbol A is location in pupil of maximum acuity, m is maximum sensitivity, + estimates location of optical axis. Arrows are the locations of maximum sensitivity along horizontal and vertical lines indicating paths of traversing test beam. The lower panel depicts movement of the Stiles-Crawford function toward the pupil center induced by wearing a contact lens with a centered artificial pupil.
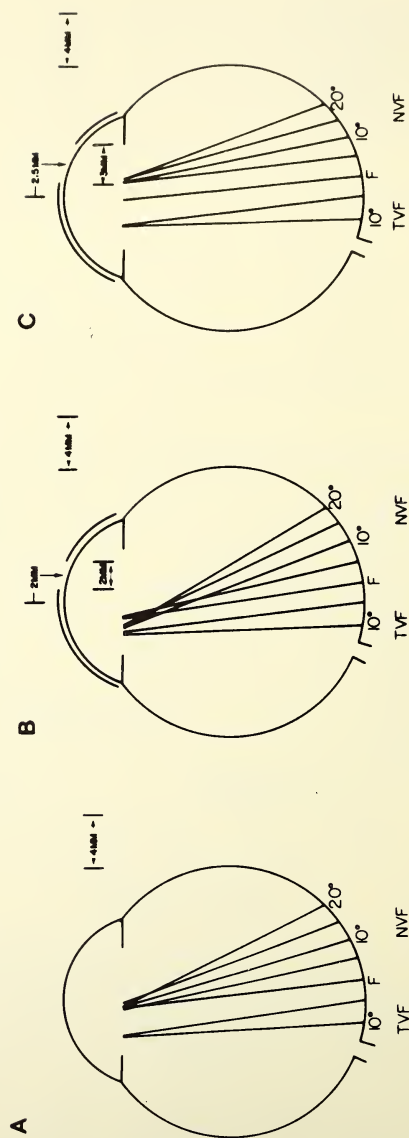
Fig 3. A schematic representation of receptor alignment characteristics in Enoch's eye (from Enoch and Birch, 1980). The lines represent axial projections of the alignment of receptors sampled at the indicated retinal loci. See text for details.

different conditions reflect a different sum, or balance of forces influencing receptor alignment. Alterations resulting from a displaced artificial pupil strengthens the arguement of a retinal origin of the S-C effect. It is clear that measurement of the S-C effect at different locations is necessary to characterize the properties of the orienting mechanisms.

The peak of the S-C effect has also been found to be displaced toward the notched (keyhole) pupil in a case of iris coloboma.[13] A phototropic mechanism is implied. Unfortunately, it is not possible to draw conclusions about the properties of an alignment mechanism from the reported single measurement at fixation. Moreover, since there was also a coloboma of the retina in this case, extending downward and nasal from the optic disc, more than one mechanism or force may have influenced alignment.

There have been reports of an apparent failure of the overall center-of-the-pupil alignment mechanism.[16,21,22,23] S-C functions measured in an amblyopic individual are illustrated in Fig. 4. Transretinal receptor alignment in the amblyopic eye is approximately toward the center of the pupil. In the nonamblyopic eye the S-C peak locations in the pupil translate opposite to the direction of the locations tested on the retina. As shown in the schematic of this eye, receptor alignment approximates a center-of-the-retinal sphere tendency. Fig. 5 summarizes similar results obtained from an eye with congenital aniridia.[24] The axial projections of the peak of the S-C function cross proximal to the center of the retinal sphere.[24] No retinal abnormalities were apparent from ophthalmoscopic examination of these amblyopic and aniridic individuals.

The results in these cases depart significantly from the usual center-of-the-pupil orientation characteristic. Some other force, or forces establish the S-C function. The orienting force is non-phototropic and is expressed approximately perpendicularly to the retinal surface, perhaps the result of active adhesion of the retina to the pigment epithelium. This center-of-the-retinal sphere orienting force is normally dominated by a phototropic (pupil-orienting) force. When the phototropic force is weak or absent, the other force defines the S-C function. Despite an apparent failure of a phototropic orienting mechanism, receptor orientation in the central retina is quite good. This is perhaps not surprising considering the geometry of the eyeball. In addition, the close packing of foveal receptors would presumably allow little dispersion of receptor orientation, contributing to the group directional sensitivity. Heath and Walraven[1] demonstrated from their measurements of the S-C effect over a distance of about 2 degrees in the central retina that the receptors are parallel to each other (Fig. 6). Packing combined with a retinal adhesion force in the central retina would act to maintain pupillary alignment in the absence of a phototropic orienting force. This may explain why prolonged total light exclusion and absence of tractional effects does not disrupt photoreceptor alignment at the center of the fovea.[25]

James E. Bailey
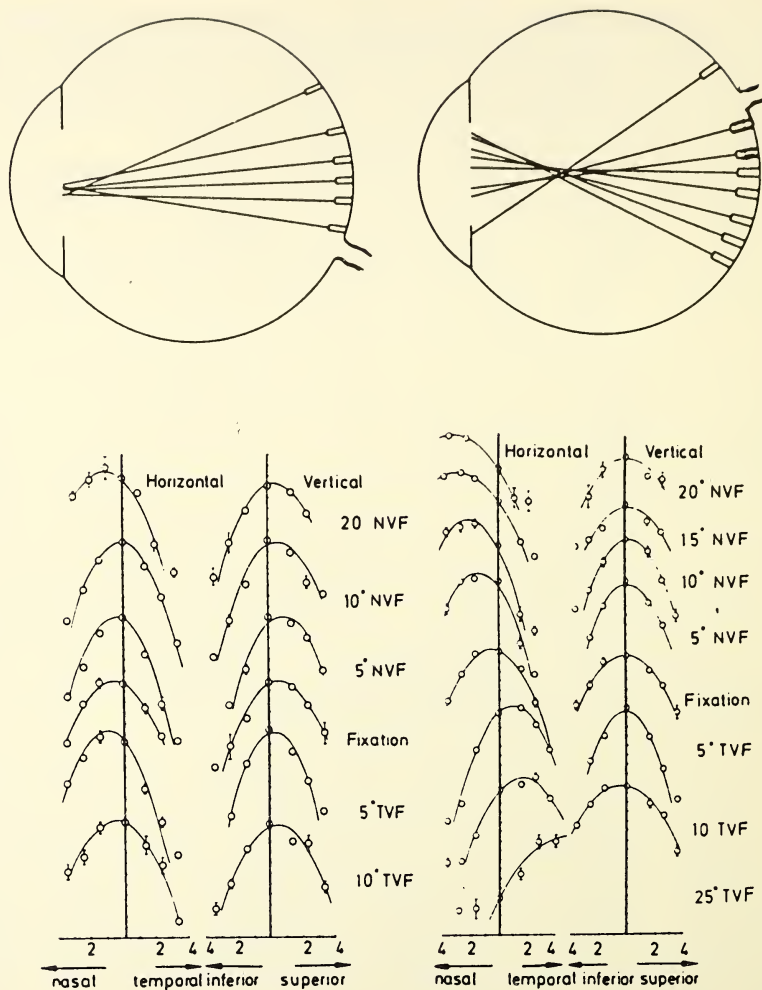
Fig 4.   Representation of receptor alignment as in Fig. 3 of an amblyopic observer (from Bedell and Enoch, 1980). The amblyopic eye and its Stiles- Crawford functions are on the left, the nonamblyopic eye on the right. Near center of the retinal sphere alignment in the nonamblyopic eye is inferred from movement of the Stiles-Crawford function opposite to the direction of movement of the test location on the retina.
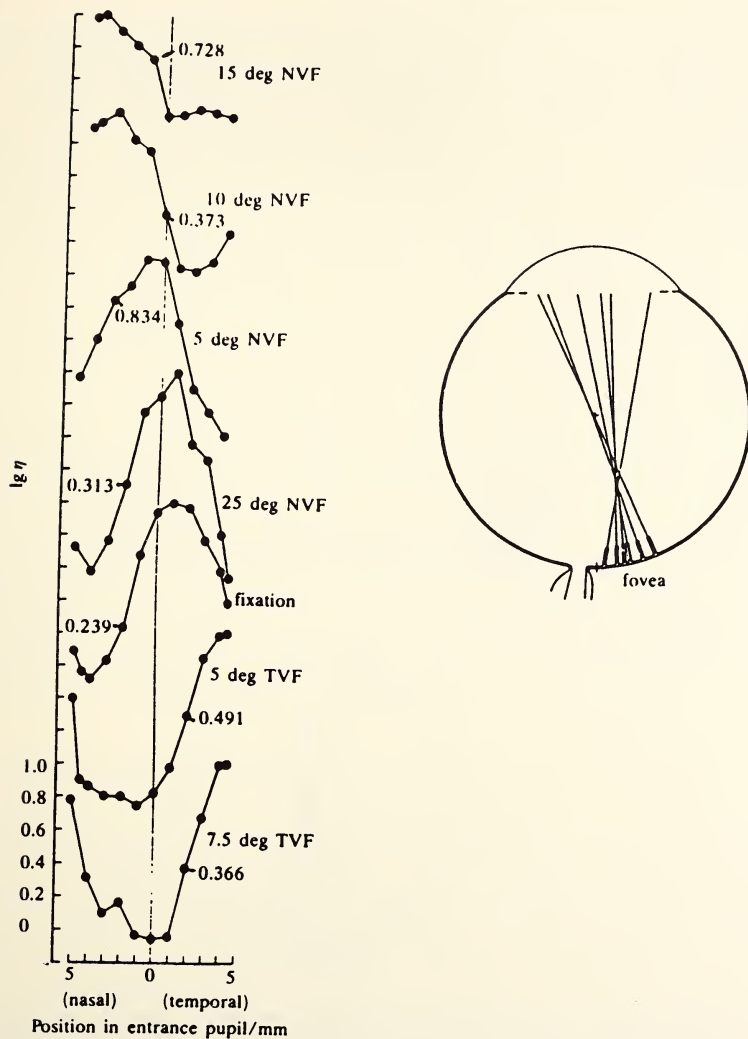
Fig 5. Stiles-Crawford functions obtained for horizontal traverses of the pupil at various locations across the horizontal meridian of the retina of an aniridic observer (from Enoch et al., 1986). The eye schematic represents projections of the peak locations to the pupil plane to indicate near center of the retinal sphere receptor alignment.
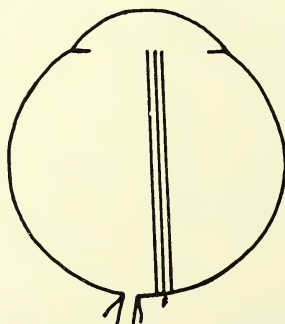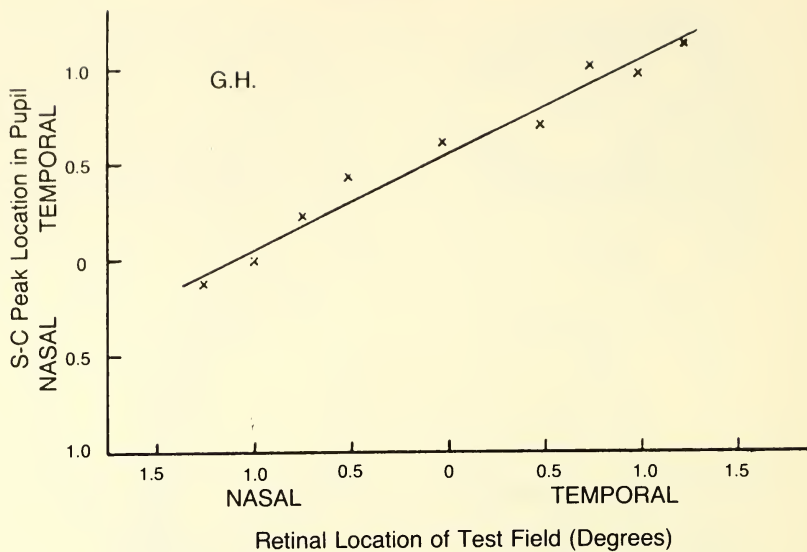
Fig 6. Location of the peak of the Stiles-Crawford function in Heath's eye (Heath and Walraven, 1970) as a function of the retinal test locus in the central retina. The change in the S-C peak location is approximately equal to that of the retinal location, indicating parallel orientations of receptors as shown by the projections in the eye schematic.

At other sites toward the retinal periphery an additional force (phototropism) is needed to orient the receptors within the boundaries of the pupil.

A working hypothesis has emerged from consideration of the foregoing experimental results.[24] The S-C effect is a measure of a (vector) sum, or perhaps a kind of balance of several forces acting to influence receptor orientation. A prenatal set-up mechanism establishes anterior pointing before birth.[26] Orientation is maintained by a pupil-orienting (assumed phototropic) force, a force perpendicular to the retina orienting receptors toward the center of the retinal sphere, tractional forces influencing local alignment, and a packing factor acting on small groups of receptors. Phototropism ordinarily dominates, but when weakened, suspended, or overcome by other factors the resultant orientation force vector may be altered. The changes may be local, when tractional forces dominate, or general due to a greater influence of an orienting force perpendicular to the retina. Nevertheless, it is clear that one must test several different sites in an attempt to characterize which force or forces defines orientation. In what follows results will be presented of recent studies of the S-C effect in congenital colobomas of the pupil and retina. These conditions offer a means to examine a possible dominance of one orienting force over another. Anomalous S-C functions were obtained in all cases studied, and the results suggest that mechanical tractional forces affect receptor orientation in the retina. Details will appear in other publications.

Stiles-Crawford Effect in Iris and Retinal Coloboma

For some time I have been seeking an observer with a simple coloboma of the iris. According to Duke-Elder,[27] this condition, known as iridoschisma, occurs in the absence of a coloboma of the retina. In Dunnewold's[13] observer there was a typical coloboma of the uvea; that is, both retina and iris were involved. This fact has been overlooked in attributing the downward 1.5 mm. displacement of the peak of the S-C effect, measured at fixation, solely to a corresponding displacement of the pupil center due to coloboma (i.e., phototropism). The coloboma of the retina was also directed downward and in (nasal) from the inferior margin of the optic disk. Furthermore, the S-C effect was not tested at different sites to rule out traction effects. Therefore, we cannot be sure that phototropism is the predominant force determining receptor alignment in this case.

An observer with congenital iridoschisma in one eye was located, and the S-C effect measured extensively. The defect in the affected left eye of this 47-year-old white male is total and complete—the whole thickness of a sector of the iris to the ciliary border is absent. This notch or cleft is located inferior and nasal at the 7 o'clock position, similar to Dunnewold's

observer. The retina of the affected eye appears normal by ophthalmoscopic examination, the ocular media are clear, and Snellen visual acuity is 20/25.

Experimental Methods

An increment threshold technique is used to measure the S-C function. It is a modification of Stiles' "field sensitivity" technique.[15] A flashing spot subtending 1 deg. is seen superimposed on a 5 deg. background. The pupil entry point of the test spot is fixed near the pupil center, that of the background is varied across the pupil. For each entry position of the background field the observer determines the luminance of the test spot necessary for threshold disappearance. Changes in the brightness of the background due to different pupil entry points affect the threshold of the test spot. Hence, it is necessary to determine in a separate test the luminance of the background required to produce the spot increment thresholds obtained in the S-C measurement. Since Weber's law is valid under these conditions ($\Delta L/L$ = constant), logarithmic changes in test field intensity are equivalent to equal logarithmic changes in the background. The "Weber range" must therefore be determined for each observer when using this indirect method of measurement of the S-C function.

Results

Horizontal and vertical S-C function determined at the point of fixation of the affected left and normal right eyes are shown in Fig. 7. Pupil position was constantly monitored with an infrared imaging system. Fixation was excellent in both eyes. The horizontal S-C functions of the two eyes are nearly identical, and the peaks are within 1 mm. of the pupil centers. Measurement of the S-C effect along a vertical traverse of the pupil produced different results. Although the S-C function of the normal, right eye is approximately centered, that of the affected left eye is outside the pupil, estimated 4 to 5 mm. superior to the center. Fig. 8 depicts results of measurements at test field locations ranging from 2.25 deg. superior to 7.0 deg. inferior in the visual field. The S-C functions at these locations are virtually identical to that obtained at fixation. Similar results were obtained at a number of additional sites in the central retina. Clearly, these functions do not peak about the center of the anomalous pupil. Instead, the inferred alignment of receptors is toward a point located somewhere beyond the superior pupil margin, opposite the direction of the downwardly displaced pupil center. There is a suggestion from these results of a dominant tractional or mechanical force drawing receptor alignment beyond the pupil boundary. However, lack of differential expression at the different retinal sites tested precludes definition of the presumed force's characteristics.

In view of the possibility of a mechanical traction force influencing receptor orientation in coloboma, additional observers were sought. As part of this study it was decided to also investigate individuals with a coloboma of the retina but without pupillary involvement. Dr. Karen Walker-Brandreth of the UC Berkeley School of Optometry suggested study of the S-C effect in individuals with Fuch's coloboma. Fuch's coloboma is a congenital anomaly of the retina associated with tilting (dysversion) of the optic nerve head. There is also typically a thinning (ectasia) of the retina inferior to the optic disc. Riise[28] considers Fuch's coloboma to be an incomplete coloboma of the retina. The iris is normal. This condition has also been studied by another contributor to this symposium.[29]

Francois[30] believes that all coloboma defects represent the same fundamental lesion, originating from the same genotypic factor but varying in penetrance and expressivity. Of additional interest is a note published by Kommerrell[31] who points out that the optic disc dysversion of Fuch's coloboma could cause a tilting of cones, altering the S-C effect and perhaps also reducing visual acuity. To date, three observers with Fuch's coloboma have been studied and the results are presented in the next section. In all cases disruption of photoreceptor alignment is apparent: marked reduction (flattening) of the S-C effect and displacement of the peak of the S-C function beyond the pupil boundary.

Figs. 9 and 10 are horizontal and vertical S-C effect functions of a Fuch's coloboma observer, an 18-year old white female with Snellen 20/40 visual acuity and excellent fixation in the measured left eye. Clearly, both S-C functions are quite flat. A flattened S-C function could result from either an increase in the acceptance angle of receptors or splaying of their alignment. The selective adaptation technique[11] was used to decide between these two possibilities. This testing situation revealed an apparent splaying.

The results of conventional S-C function measurements at other retinal locations are compared with those obtained at fixation in Fig. 11. Horizontal functions remain flat as the test site moves toward the optic disc (temporal visual field) and steepen somewhat beyond 5 deg. in the nasal field. The vertical S-C functions are more variable. Changes in both shape and peak location are evident, and the functions steepen as the test location is moved away from the optic disc and ectasia region. Clearly, the forces influencing receptor alignment are complex in this eye. There is no indication of a limited site of origin. Perimetric testing (Goldmann I 3e) of this eye revealed a relative scotoma in the superior temporal visual field. It could be eliminated by adding a -4.0 diopter lens to the observer's spectacle correction, resulting in the now overcorrected, surrounding "normal" visual field becoming scotomatous. The location of this refractive scotoma corresponded to the central region of the retinal ectasia (inferior to the optic disc). These
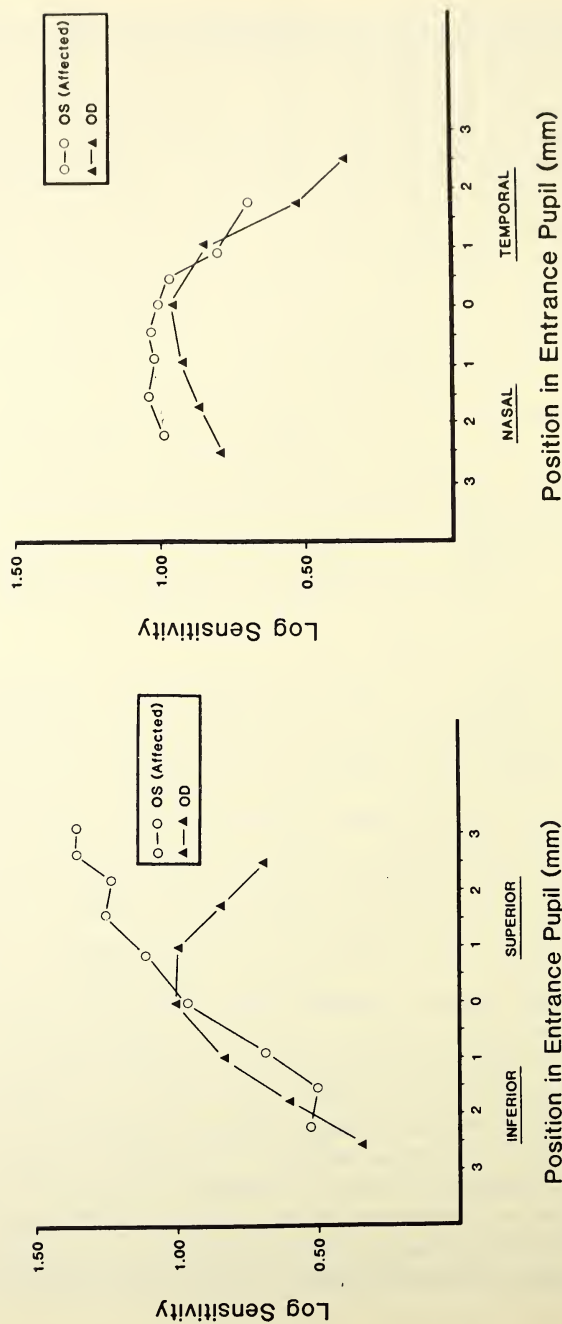
Fig 7. Stiles-Crawford effect data of the affected and nonaffected eyes of an observer with simple coloboma of the iris. Vertical traverses are shown in the left panel, horizontal in the right. The 'zero' position in the entrance pupil of the colobomatous eye refers to the 'center of gravity' of the pupil.
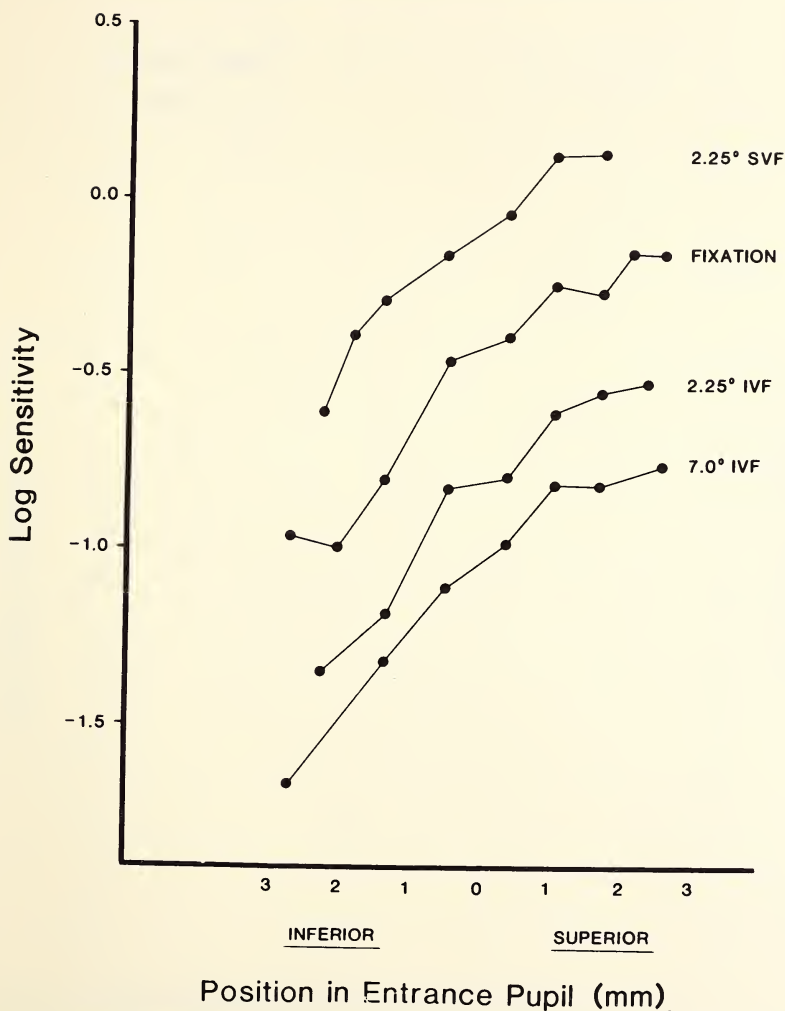
Fig 8. S-C effect data obtained for vertical traverses across a pupil with a coloboma at different visual field locations. Testing extended into the notched area of the pupil inferior to center. The functions are displaced vertically for clarity of presentation.
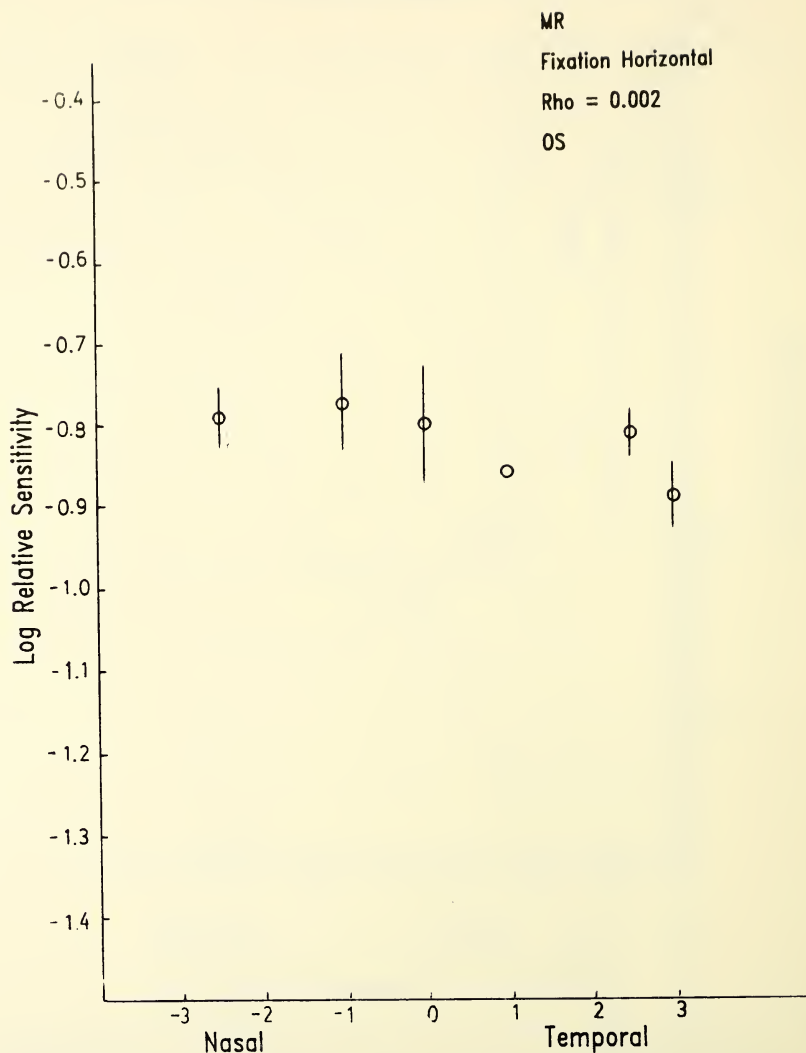
Fig 9. Stiles-Crawford data obtained for a horizontal traverse covering approximately 6 mm. across the pupil of an observer with Fuch's coloboma. Testing was at the point of fixation. The rho value indicates the relative steepness or flatness of the function. Normally it is about 0.05. The lower the value the flatter the function. The vertical bars represent the range of 1 standard deviation.
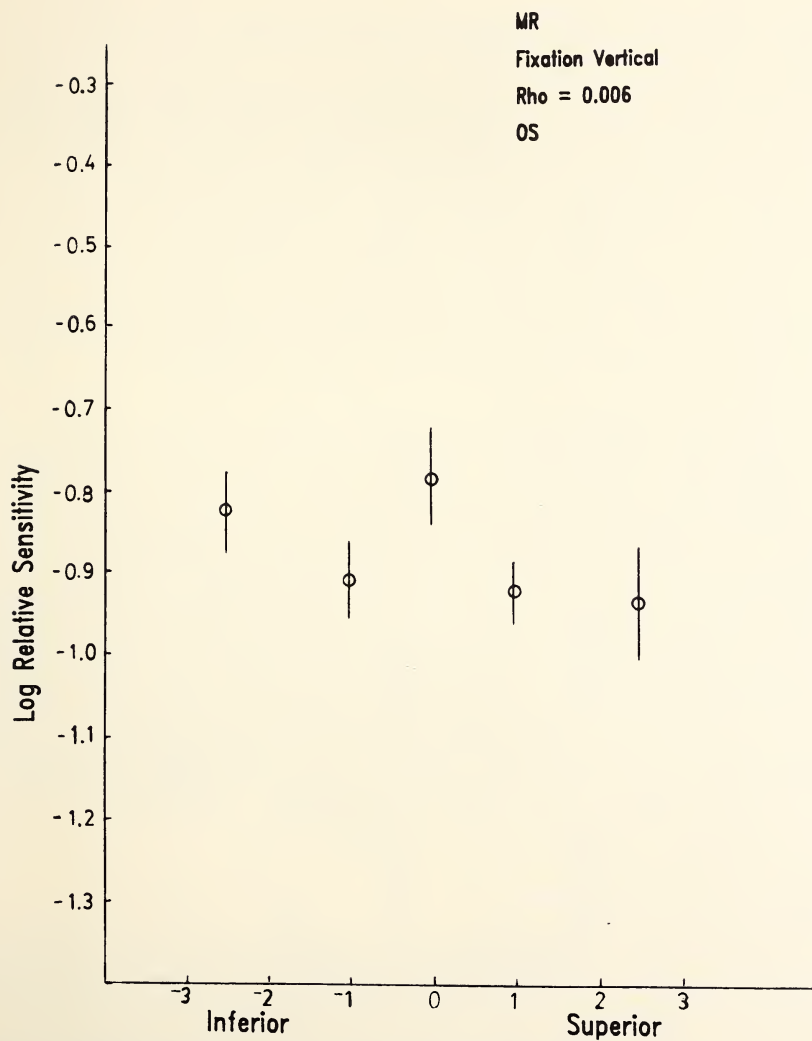
Fig 10. Same as in Fig. 9 except a vertical traverse across the pupil.

findings indicated a more myopic (extended) position of the ectasia area relative to the surrounding retina.[32]

Fig. 12 depicts similar results, though not as extensive, in another Fuch's coloboma observer, a 38-year old black female with Snellen 20/30 visual acuity and a relative scotoma extending in a direction superior and temporal from the blind spot (inferior and nasal from the optic disc). At fixation the vertical S-C function is flat. The horizontal measurements indicate that the S-C function peaks outside the pupil. Both functions are flat at the temporal field location nearest the optic disc and ectasia. The vertical function steepens as the test location is moved into the nasal visual field beyond fixation but the horizontal function remains grossly displaced nasalward in the pupil. Again, the forces influencing alignment are complex in this eye.

Results of S-C effect measurements in a third observer with Fuch's coloboma at a number of different retinal locations are shown in Fig. 13. This individual is a 21-year old white male with Snellen 20/25 visual acuity in the affected right eye. The horizontal S-C function obtained at fixation appears essentially normal; the peak of the vertical function is displaced outside the pupil in the inferior direction. At test locations beyond fixation in the nasal visual field the vertical functions shifts back toward the pupil center, while at test sites in the opposite direction it moves farther beyond the inferior pupil margin. There is a graded shift of the peak of the vertical S-C function in an inferior direction as the test site is moved toward the location of the ectasia. Except for the 5.25 deg superior and temporal visual field location the horizontal S-C function remained centered in the pupil. These results suggest a more limited site of origin of a mechanical traction force, acting in the direction of the retinal ectasia.

Discussion and Conclusions

Receptor alignment was found to be disturbed in all eyes with coloboma in this study. Although phototropism usually is the dominant force considered to influence overall receptor alignment, clearly it is not in these coloboma defects. This is not to say that a phototropic alignment mechanism is completely non-functional. However, the results suggest some other retinal components—a variety of tractional effects—are influencing the results obtained. These effects are not visible by ophthalmoscopy, particularly in the case of iris coloboma. Surprisingly, the peak of the S-C effect function was displaced opposite to the direction expected from the action of a phototropic alignment mechanism. This mechanism would serve to orient the receptors more toward the location of the notch in the pupil. Phototropism, although possibly still active, appears to be dominated by a mechanical traction force. The nature of this force is not clear due to lack

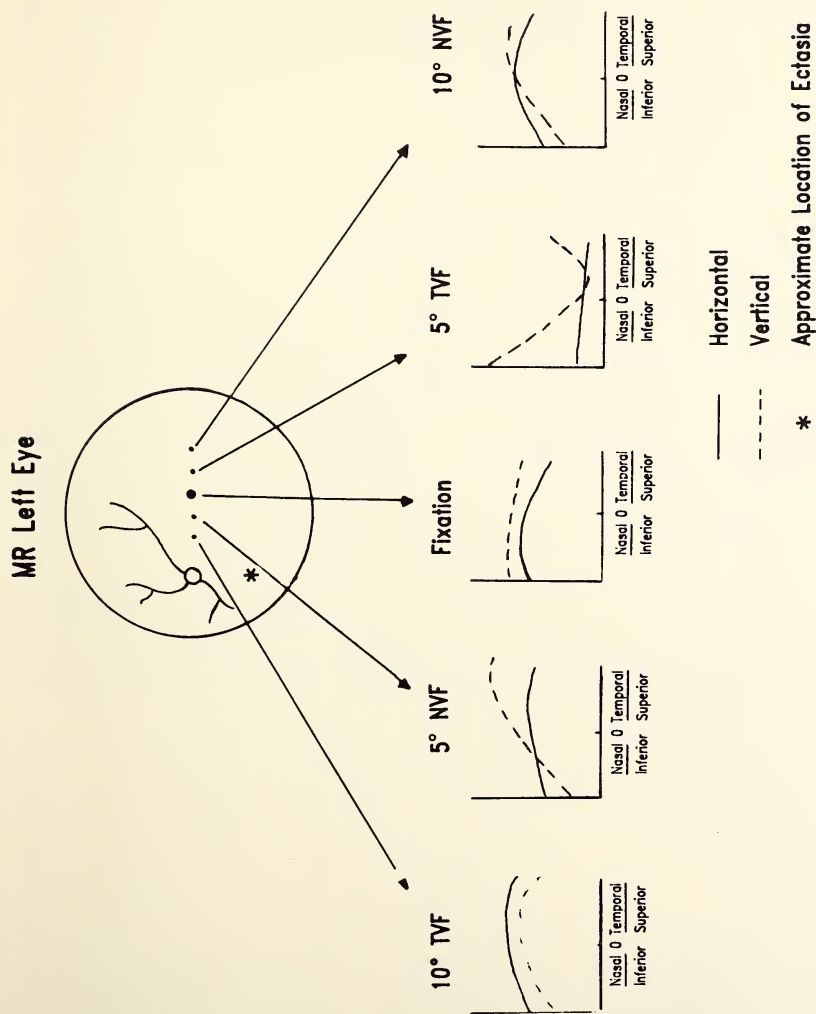Fig 11. A summary of Stiles-Crawford functions obtained for horizontal and vertical traverses across the pupil showing the different visual field locations tested in the observer whose S-C effect data at fixation is shown in Figs 9 and 10. NVF (TVF) refers to nasal (temporal) visual field. The (*) marks the approximate location of the center of an ectasic retinal area.
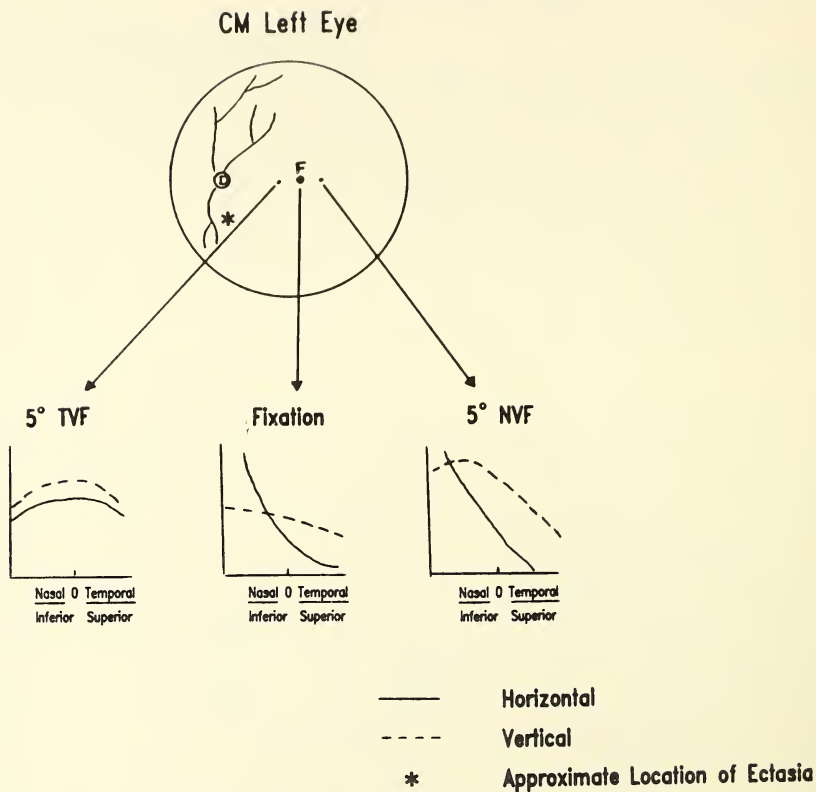
James E. Bailey

Fig 12. A summary of S-C effect data like that presented in Fig 11 but obtained from an additional observer with Fuch's coloboma.
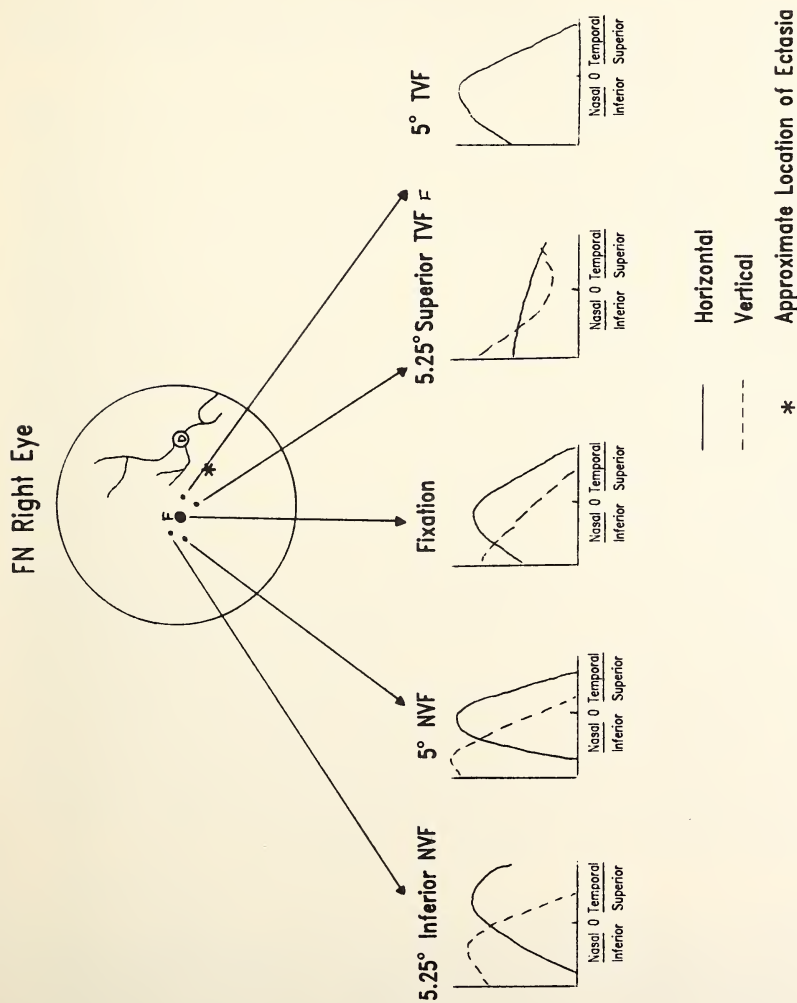
Fig 13.   A summary of additional S-C data obtained from a third observer with Fuch's coloboma.

of differential effects on the S-C function at different retinal test sites. There may be an anomaly of the morphology of the central retina resulting in a local flattening or tilting, causing receptors to be aligned outside the pupil. The close packing of receptors in this region would presumably not allow a phototropic force component to compensate for the misalignment. To date, we have not fit an artificial pupil contact lens to check for an active alignment mechanism[20,21] or conducted further studies at additional retinal sites to rule out phototropic function.

The suggestion of mechanical traction forces due to ectasia and malinsertion of the optic nerve in Fuch's coloboma led to a study of the S-C effect in this condition. Evidence of traction expressed in a variety of forms influenced alignment in all observers. In one observer splaying of receptor alignment was apparent from selective adaptation testing. Conceivably, this is the result of the effect of force vectors with different directions of action dispersing receptor alignment.[24] Testing at other sites revealed gross tilting of receptors, attributed to a limited origin or a complex set of tractional effects. Mechanisms influencing alignment are apparently somewhat functional at some sites but it is not clear whether they are totally normal. Clearly, however, phototropism is not fully effective in compensating for the influence of other alignment forces.

Perhaps a general characteristic of coloboma is an anomaly of the prenatal set-up mechanism subserving receptor alignment. The commonality of coloboma defects has been pointed out.[30] If phototropism is precluded from functioning, traction associated with the condition dominates. The presence of a retinal coloboma in Dunnewold's[13] case of iris coloboma means traction cannot be ruled out as an influence on alignment. Furthermore, the results of this study serve to underscore the fact that one measure of the S-C effect does not allow a complete analysis of mechanisms defining orientation. Additional measures aid in characterizing the pathophysiology of the condition under study. This is the essence of diagnosis. The Stiles- Crawford effect is a useful tool for evaluating underlying pathophysiological effects influencing receptor alignment.

References

1. Heath GG Walraven PL. Receptor orientations in the central retina. J Opt Soc Am 1970; 60:733-4
2. Heath GG. Directional sensitivity of retinal receptors. J Opt Soc Am 1970; 60: 736.
3. Bailey JE Heath GG. Flicker effects of receptor directional sensitivity. Am J Optom Physiol Opt 1978; 55(12): 807-12.
4. Bailey JE. Magnitude variation of the Stiles-Crawford effect in the central retina. ARVO Suppl 1975; 102.

5. Bailey JE. Influence of the crystalline lens on the Stiles-Crawford effect. J Opt Soc Am 1982; 72: 1758.
6. Stiles WS Crawford BH. The luminous efficiency of rays entering the eye pupil at different points. Proc Roy Soc London (B) 1933; 112: 428-50.
7. Crawford BH. The luminous efficiency of light entering the eye pupil at different points and its relation to brightness threshold measurements. Proc Roy Soc London (B) 1937; 124: 81-96.
8. Enoch JM Tobey FL (eds.). Vertebrate Photoreceptor Optics. Springer Series in Optical Sciences, vol. 23. Berlin: Springer 1981, pp. 144-9.
9. Makous WL. A transient Stiles-Crawford effect. Vision Res 1968; 8: 1271-84.
10. Coble JR Rushton WAH. Stiles-Crawford effect and the bleaching of cone pigments. J Physiol 1972; 217: 231-42.
11. MacLeod DIH. Directionally selective light adaptation: A visual consequence of receptor disarray? Vision Res 1974; 14: 369-78.
12. Weale RA. On the problem of the retinal directional sensitivity. Proc Roy Soc London 1981; 212(B): 113-30.
13. Dunnewold CJW. On the Campbell and Stiles Crawford Effects and Their Clinical Importance. Soesterberg: The Institute for Perception TNO 1964.
14. Applegate RA Meade DL Christina M. Normal variation in the Stiles-Crawford function peak location. Non-Invasive Assessment of the Visual System Technical Digest, 87-4 (Opt Soc Am) Washington DC 1987, pp. 15-18.
15. Enoch JM Hope GM. An analysis of retinal receptor orientation III. Results of initial psychophysical tests. Invest Ophthalmol Vis Sci 1972; 11: 765-82.
16. Bedell HE. Central and peripheral retinal photoreceptor orientation in amblyopic eyes as assessed by the psychophysical Stiles-Crawford function. Invest Ophthalmol Vis Sci 1980; 19(1): 49-59.
17. Bedell HE Enoch JM. A study of the Stiles-Crawford (S-C) function at 35° in the temporal field and the stability of the foveal S-C function peak over time. J Opt Soc Am 1979; 69: 435-42.
18. Enoch JM. Retinal receptor orientation and the role of fiber optics in vision. Am J Optom Arch Am Acad Optom 1972; 49(6): 455-70.
19. Bonds AB MacLeod DIA. A displaced Stiles-Crawford effect associated with an eccentric pupil. Invest Ophthalmol Vis Sci 1978; 17: 754-61.
20. Applegate RA Bonds AB. Induced movement of receptor alignment toward a new pupillary aperture. Invest Ophthalmol Vis Sci 1981; 21(6): 869-73.
21. Enoch JM Birch DG. Evidence for alteration in photoreceptor orientation. Ophthalmology 1980; 87(8): 821-33.
22. Enoch JM Birch DG. Inferred positive phototropic activity in human photoreceptors. Phil Trans Roy Soc London 1981; 291B: 323-51.
23. Bedell HE Enoch JM. An apparent failure of the photoreceptor alignment mechanism in a human observer. Arch Ophthalmol 1980; 98: 2023-26.
24. Enoch JM Lakshminarayanan V Yamade S. The Stiles-Crawford effect of the first kind (SCEI): Studies of SCEI in an aniridic observer. Perception 1986; 15: 777-84.
25. Applegate RA Adams AJ Bradley A Eisner A. Total occlusion does not disrupt photoreceptor alignment. Invest Ophthalmol Vis Sci 1986; 27(3): 441-3.

26.   Laties AM Enoch JM. An analysis of retinal receptor orientation. I. Angular relationships of neighboring photoreceptors. Invest Ophthalmol Vis Sci 1971; 10: 69-77.

27.   Duke-Elder S. Systems of Ophthalmology, vol. III. Normal and Abnormal Development: Part 2, Congenital Deformities. St. Louis: Mosby 1963; p. 573.

28.   Riise D. The Nasal Fundus Ectasia. Acta Ophthalmol (Suppl) 126, 1975.

29.   Williams TD. Congenital malformations of the optic nerve head. Am J Optom Physiol Opt 1979; 55(10): 706-18.

30.   Francois J. Heredity in Ophthalmology. St Louis: Mosby 1961, pp 145-85.

31.   Kommerrell G. Binasale refraktionsskotome. Klin Mbl Augenheilkunde 1969; 154: 85-8.

32.   Fankhauser F Enoch JM. The effects of blur upon perimetric thresholds. Arch Ophthalmol 1962; 68: 120-31.

THE STILES-CRAWFORD EFFECT IN RETINAL DISEASE: INTERPRETATION IN THE PRESENCE OF CATARACT
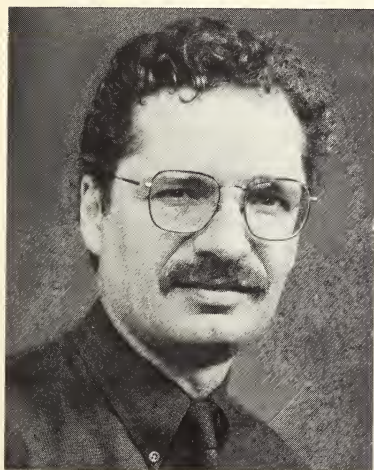
Raymond A. Applegate
Associate Professor in the Department of Ophthalmology at University of Texas Health Science Center, San Antonio, Texas. He is Feature Co-editor of the Applied Optics Annual Issue on Non-Invasive Assessment of the Visual System.

ABSTRACT. This paper briefly reviews the evidence for photoreceptor directionality, the use of the Stiles-Crawford effect (SCE) in the assessment of human photoreceptor directionality, our current understanding of human photoreceptor directionality, and the use of the SCE in monitoring, detecting and predicting impending retinal disease. Using these brief reviews as a backdrop, the paper introduces the problem of interpreting the SCE in the presence of cataracts, presents a new model of SCE behavior in the presence of cataracts, and uses the model to predict the density, location and spread of the cataract as well as the directional properties of the photoreceptor under test. Two future studies are proposed and the limitations of the new analysis technique are discussed.

THE STILES-CRAWFORD EFFECT IN RETINAL DISEASE: INTERPRETATION IN THE PRESENCE OF CATARACT

Photoreceptor Directionality

Photoreceptors have several cellular and group structural properties which cause them to be directionally sensitive to light. At the cellular level these include: Photopigment molecules within the receptor which are aligned to optimize photon capture along the long axes of the cell;[1] cell walls configured such that the cross sectional diameter of the photoreceptor is small (approximately 2 microns for foveal cones) compared to its length (approximately 60 microns);[2] and, a refractive index difference across the cell wall such that the cell's internal refractive index is greater than the refractive index of the external supporting matrix.[3-5] These cellular attributes combine to make each photoreceptor a small waveguide which funnels light down its long axis until it is either absorbed or passes into the pigment epithelium. As a group, photoreceptors across the retina align their long axes toward the center of the eye's exit pupil.[6] Together these cellular and group structural properties serve to emphasize relevant visual information (light entering through the eye's pupil) and minimize the influence of stray light.

Noninvasive Assessment of Photoreceptor Directionality

Properties of human photoreceptor alignment can be investigated non-invasively by measuring the Stiles-Crawford effect (SCE), a psychophysical determination of the relative efficiency of light in eliciting a visual response as a function of pupil entry.[7] Typically SCEs are graphically displayed by plotting log relative sensitivity (LRS) as a function of pupil entry (solid dots in Figure 1) for both a horizontal (panel A) and a vertical (panel B) pupil traverse. Because the SCE is in part attributable to specialized organization of receptor alignment and structure, the SCE can be used to infer both the principal orientational properties (peak of the SCE) and directionality (width of the effect) of the population of photoreceptors under test.[8-10]
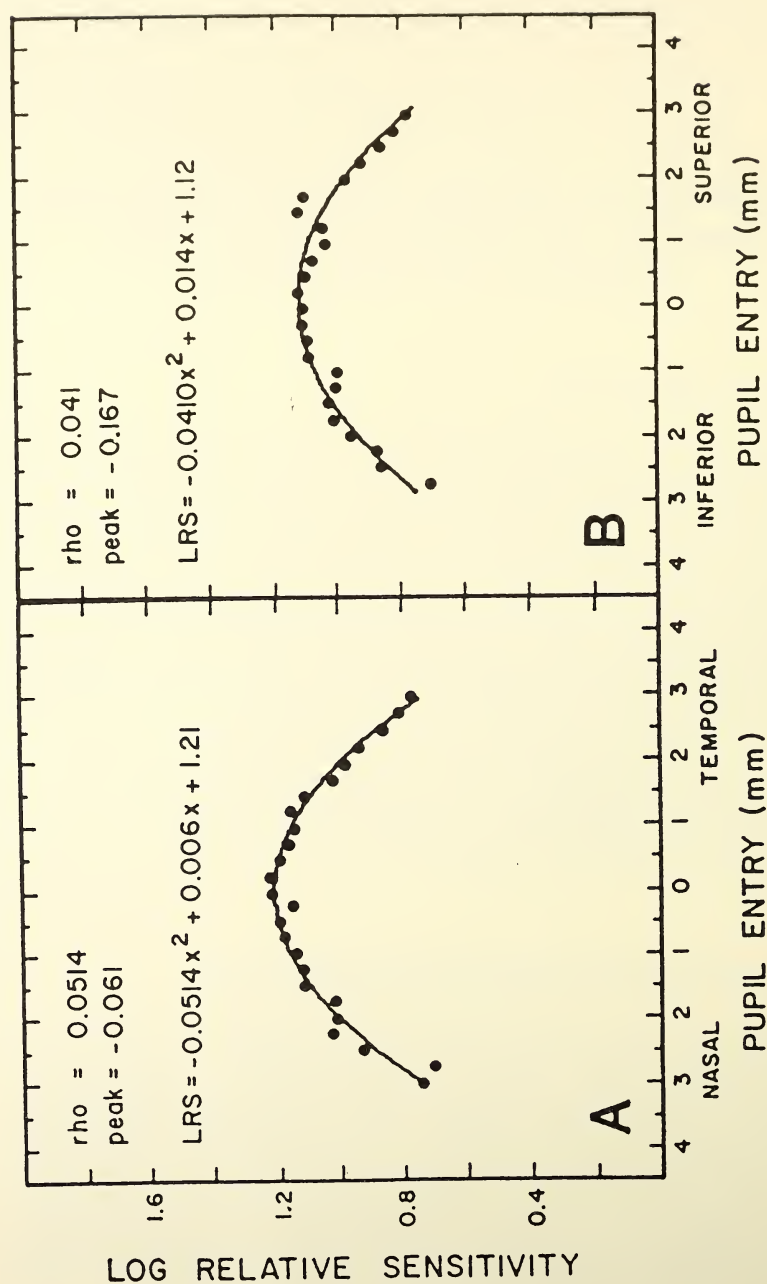
Figure 1. A normal SCE data set (solid dots) and associated SCF (smooth curve) graphically displayed as log relative sensitivity as a function pupil entry for both a horizontal (Panel A) and vertical (Panel B) pupil traverse.

In practice, a best fit parabolic curve (smooth curves Figure 1) is used to characterize each SCE data set. This best fit curve is termed the Stiles-Crawford function (SCF) and it is the calculated coefficients of the SCF that serve to define objectively the principal alignment tendencies (SCF peak location) and distributive properties (SCF rho value) of the population of photoreceptors under test (Figure 1 insets).

Current Understanding of Human Photoreceptor Directionality and the Use of the SCE in Monitoring, Detecting, and Predicting Impending Retinal Disease

Previous studies indicate that the basic pattern of receptor alignment is present at birth,[6] stable over time,[11] recovers after normal[12] and pathological stress[13] (also see below), is actively influenced by changes in the position of the pupillary aperture[14,15] and is not influenced by prolonged (6 days) total occlusion.[16-18] Combined, these studies suggest that receptor alignment is genetically programmed and actively maintained postnatally by a positive phototropic mechanism which is not disrupted by prolonged periods of total occlusion.

Together these factors combine to make the SCE an excellent tool for monitoring the local integrity of the photoreceptors in cases of retinal pathology. This potential has not gone unnoticed. Fankhauser et al were the first to document alterations in the SCE associated with retinal pathology.[19] Of particular importance in this study was the indication of a possible recovery of the SCE in a case of retinal degeneration treated with photocoagulation.

Definite recovery and/or progressive alteration of the SCE has been recorded in cases of retinal detachment,[20,21] subretinal fluid accumulation,[13] acute posterior multifocal placoid pigment epitheliopathy,[22] central serous choroidopathy,[23,24] macular degeneration,[24,25] fibrovascular scar,[24] choroidal atrophy,[26] presumed histoplasmosis,[24] retinitis pigmentosa,[27] choroideremia,[24] and vitelliruptive macular dystrophy.[24] Further, in healed pathology[21,28,29] and healed traumatic injury[15] the SCF peak location has been observed to be stable over time albeit displaced from the normal central location.

While the SCE has been used to monitor receptor involvement in the natural history of various active retinal pathologies, it has received little attention as a tool for the early detection and/or prediction of impending retinal disease. There are two principal reasons for this neglect. First, SCE measurement techniques are time consuming; and second, the measurement is extremely local. Thus, even for the faster SCE measurement techniques[15,30] unless the retinal location of interest is known, sampling multiple retinal locations in an effort to detect or predict impending receptor

involvement in pathology is neither time nor cost effective in a clinical environment.

However, if the most likely retinal site of the pathology is known, then the SCE measurement may be a cost effective clinical tool. While *a priori* knowledge of the retinal site of the pathology may seem to be an unusually restrictive criterion, it is often easily met in many at risk patient populations.

One such patient population of particular interest consists of patients over the age of 50 with macular drusen. These patients are not only known to be at risk for developing age related maculopathy (ARM),[31-35] particularly if one eye is already affected,[34] but the site of retinal involvement (the macula) is specified. Given that ARM is the leading cause of blindness over the age of 50,[36,37] development of tests capable of early identification of which individuals will, or will not, go on to develop some form of exudative maculopathy are important not only for understanding the disease process but for initiating and monitoring therapy designed to alter the natural course of the disease.

As mentioned above, results using the SCE as a monitoring tool to follow receptor involvement in patients diagnosed as having ARM are encouraging.[24,25] This point is illustrated in Figure 2 which details previously unreported data for a 53 year old white female ARM patient with binocular drusen and a small pigment epithelial detachment in one eye and good visual acuity (20/25 exudative eye, 20/20 non-exudative eye). While these findings are encouraging, the interpretation of SCE measurements made in other at risk patients in this age group revealed a serious complicating factor—the presence of cataracts. While cataracts manifested themselves as an obvious defect in the SCE in some patients (focal dense cataract) in others the effect on the SCE was large and often well concealed (senile cataract).

The Problem of Cataracts

While cataracts can manifest themselves in a myriad of ways, Figure 3 illustrates how even a simple well-behaved cataract can markedly influence the psychophysically measured SCE without markedly influencing its parabolic nature. Panel A of Figure 3 illustrates an idealized SCE in the absence of any cataract (i.e. the calculated SCF and the measured effect are identical). The peak of the function is centered (SCF peak location $= -b/2a = 0$) and the spread or rho value is 0.05 (rho $= -a = 0.05$). Panel B models the loss in log sensitivity induced by a series of cataracts of increasing severity assuming the density characteristics of the cataract within the pupil varies in a Gaussian manner. In this model of cataract behavior, the coefficient 'A' reflects the maximum density of the cataract, 'd' the

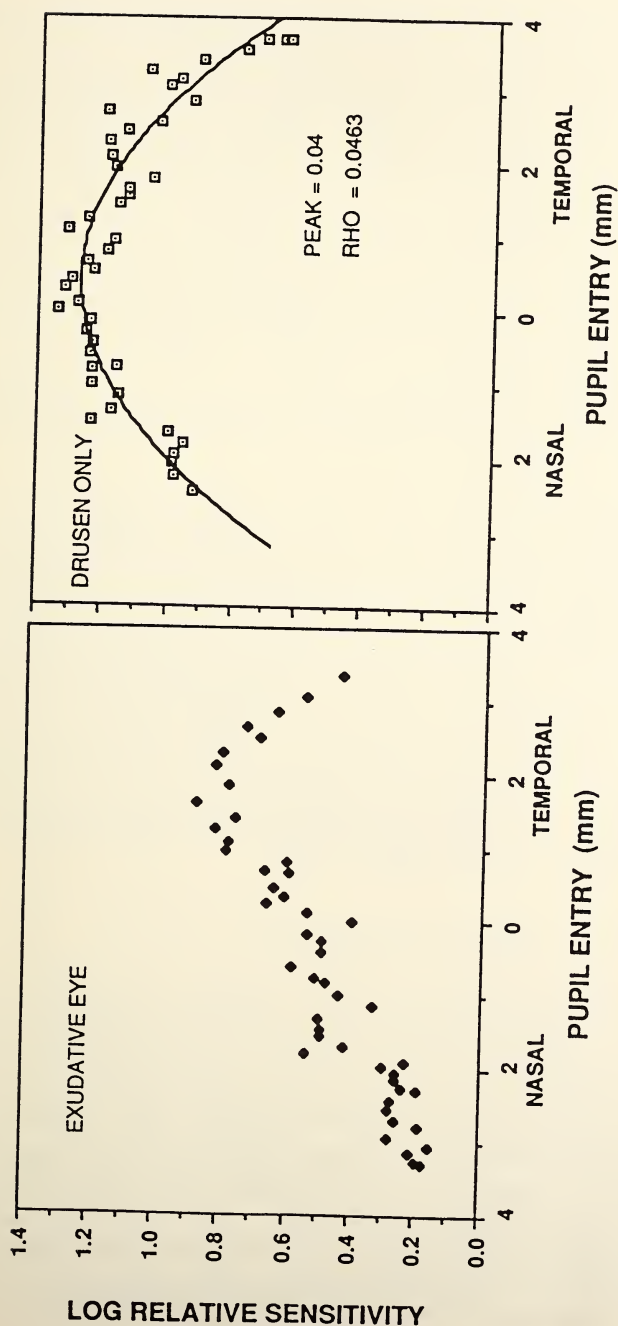**Figure 2.** SCE data sets from horizontal pupil traverses of both eyes of a 53 year old white female patient with ARM. Panel A illustrates data from the eye with good visual acuity (20/25), macular drusen and a small pigment epithelial detachment within the macular area inferior and nasal to the fovea. Panel B illustrates data for the contralateral eye having good acuity (20/20) and only macular drusen.

location of maximum cataract density and 'σ' the spread of the density effects of the cataract along the pupil meridian of interest. In the series of cataract effects modeled in panel B, maximum density (A) of the cataract varies from 0 (no cataract) to 0.4 in steps of 0.1 for each successive curve, the pupil location of maximum density (d=0) is fixed centrally (pupil entry 0), and the spread (standard deviation) of the density effects (σ) is fixed at 2.0. Panel C illustrates the expected log relative sensitivity of the measured cataract-contaminated SCE for cataract densities 0 (no cataract, top curve) to 0.4 (bottom curve) as a function of pupil entry assuming simple summation of the two components (the underlying directional properties of the receptor reflected in the SCE, panel A, and the loss in log sensitivity resulting from the density characteristics of the cataract, panel B). Notice the large effect even a mild cataract (.1 to .4) is predicted to have on the measured SCE (regular sunglasses have a uniform density of 1.0 - 2.0).

To understand the potential error induced by analyzing a SCE data set contaminated by the presence of a cataract, consider how a cataract-contaminated SCE would be interpreted using the traditional analysis procedures. For illustrative purposes, assume the SCE seen in Figure 3 panel A is measured at 7 pupil entries (1 mm intervals over a range ±3 mm from the SCF peak) in the presence of the 0.4 density cataract modeled in Panel B of Figure 3. The anticipated cataract-contaminated SCE data set would consist of 7 data points taken at 1 mm intervals from the bottom curve of Figure 3 panel C. This data set is illustrated in Figure 4 as solid dots. Fitting this cataract-contaminated SCE data set with the traditional parabolic function (smooth curve) would lead to the conclusion that there is little if any receptor directionality in this eye (rho = 0.009).

A Potential Solution to the Cataract Problem

To solve this misinterpretation problem, Applegate and Massof have proposed fitting the cataract-contaminated SCE data set with the new model of SCF behavior presented in Figure 3 Panel C using a reiterative least squares analysis.[38] The coefficients of the new model describing the best fitting function would in turn be used as estimates of the density (A), location (d) and spread (σ) of the cataract as well as the principal alignment tendencies (-b/2a) and distributive properties (-a) of the receptors under test. The model assumes the parabolic nature of the underlying retinal directionality remains in the presence of a cataract and that the loss of log sensitivity as a function of pupil entry resulting from the influence of a cataract is additive and can be represented by a negative Gaussian.
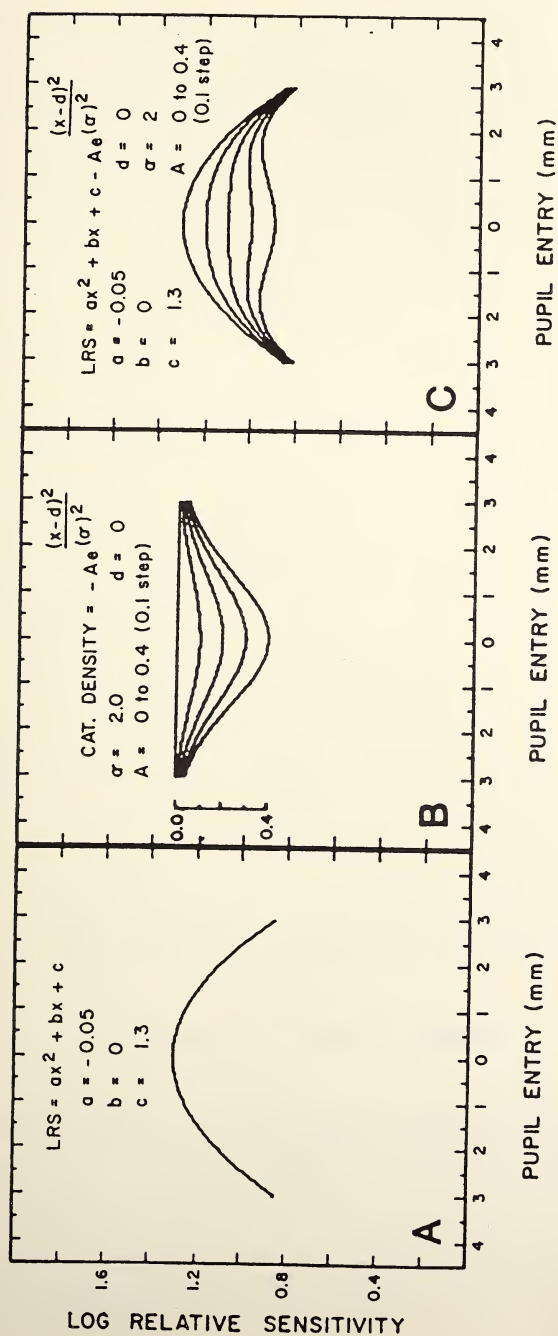
Figure 3. Anticipated effects of a series of mild diffuse cataracts on the measured SCE. Panel A displays an idealized SCE data set prior to perturbation by a cataract. Panel B illustrates the loss in log sensitivity resulting from the density properties of a series of cataracts of increasing density. Panel C illustrates the combined effects of receptor directionality (Panel A) and cataract induced loss in log sensitivity (Panel B) on the measured SCE.
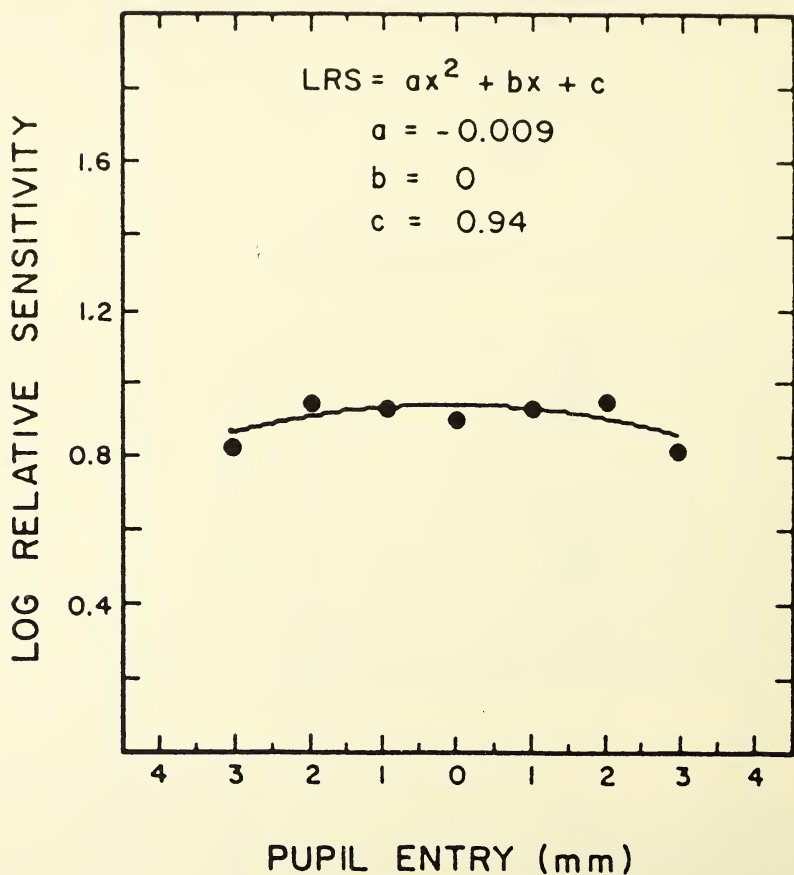
Figure 4. Traditional parabolic analysis of a cataract-contaminated SCE data set leads to erroneous estimates of inferred receptor directional properties.

Does the Model Work?

Example 1. Figure 5 graphically displays a cataract-contaminated SCE (solid dots) in a senior citizen having drusen, 20/20 acuity and mild diffuse cataract (lens yellowing). Panel A illustrates the model's best fitting curve to the underlying cataract-contaminated SCE data set. Panels B and C break the best fitting function into its two components—the influence of the cataract on log sensitivity (Panel B) and receptor directional properties (Panel C). Panel B details the model's estimates for the cataract's maximum density (A = 0.47), the location of maximum density (d = 0.05), and the spread of the density across the pupil (σ = 1.70). The solid curve illustrates the cataract-induced loss in log sensitivity as a function of pupil entry. Panel C details the model's estimates of the underlying receptor directional properties in the presence of the cataract (solid curve—parameters listed to the right) and the traditional parabolic analysis which ignores the presence of the cataract (dashed curve—parameters listed to the left). Notice that while the SCF peak location remains essentially constant between the traditional parabolic analysis and the parabolic portion of the new model, the rho value is estimated to be greater by a factor of 2 (0.043 to 0.087) using the new model.

Example 2. An obvious cataract effect can be seen in SCE measured in individuals having dense localized cataracts. Figure 6 illustrates the SCE (solid dots) for both a horizontal (Panel A) and vertical (Panel B) pupil traverse of a retinitis pigmentosa patient having a relatively dense posterior subcapsular cataract. As anticipated, these cataract-contaminated SCE data sets are not well represented by a simple parabolic analysis (solid curves). Figures 7 and 8 (Panel A) graphically displays the effect of applying the new two component model of SCE behavior to both the horizontal (Figure 7) and vertical (Figure 8) cataract-contaminated SCE data sets. Panel B and C, in these Figures show the two components of the model individually— Panel B, the negative Gaussian representing the loss in log sensitivity resulting from the estimated density effects of the cataract and Panel C the parabola representing the orientational properties of the receptors.

While the new model impressively reflects the detail of the cataract-contaminated SCE data set and is therefore an encouraging analysis technique, the improvement in fit over the traditional parabolic analysis is not particularly surprising. The new model is a more complex model with more free parameters and should *a priori* better fit the underlying data set. The test of the new model is not only how well the cataract-contaminated SCE data set is fit by the model, but whether or not the two components of the model reflect the true underlying effects of cataract and the receptor directional properties on the cataract-contaminated SCE. These questions are yet to be answered.
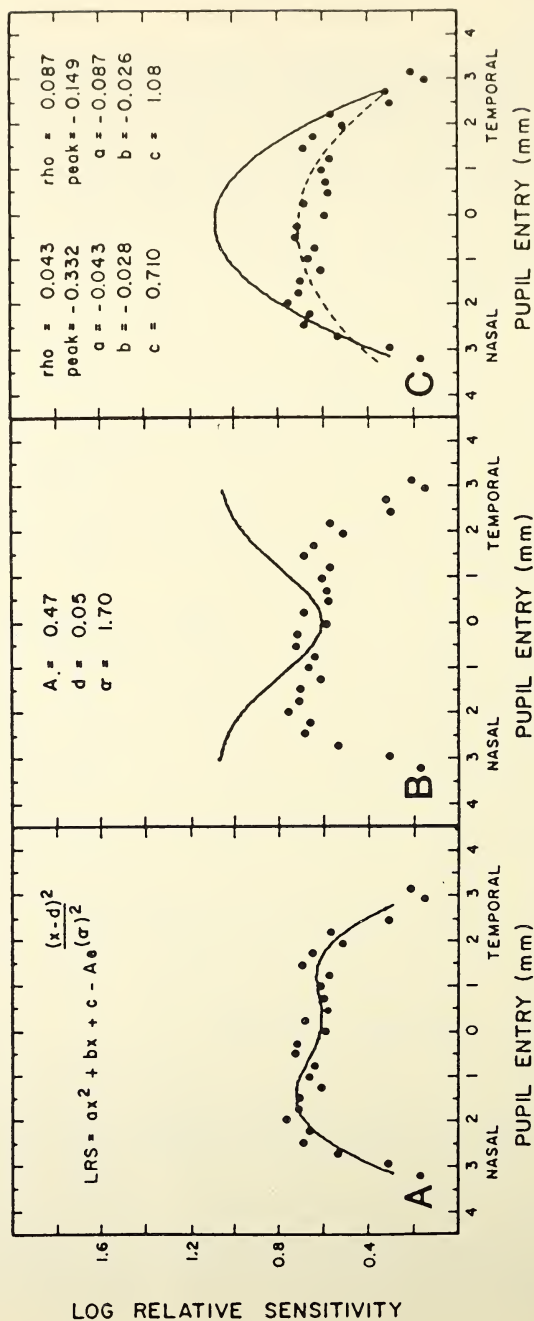
Figure 5. The new two component model of SCE behavior applied to analyze the impact of a diffuse cataract on the underlying directional properties of the photoreceptors. Panel A displays the model's best fitting function to the cataract-contaminated SCE data set. Panel B illustrates the loss in log sensitivity resulting from the model's estimate of the cataract's maximum density (A), location of maximum density (d), and spread (σ). Panel C displays both the new model's estimate (solid curve - parameters of parabola listed on the right) and the traditional model's estimate (dashed curve - parameters of parabola listed on the left) of parameters of the best fitting parabola.
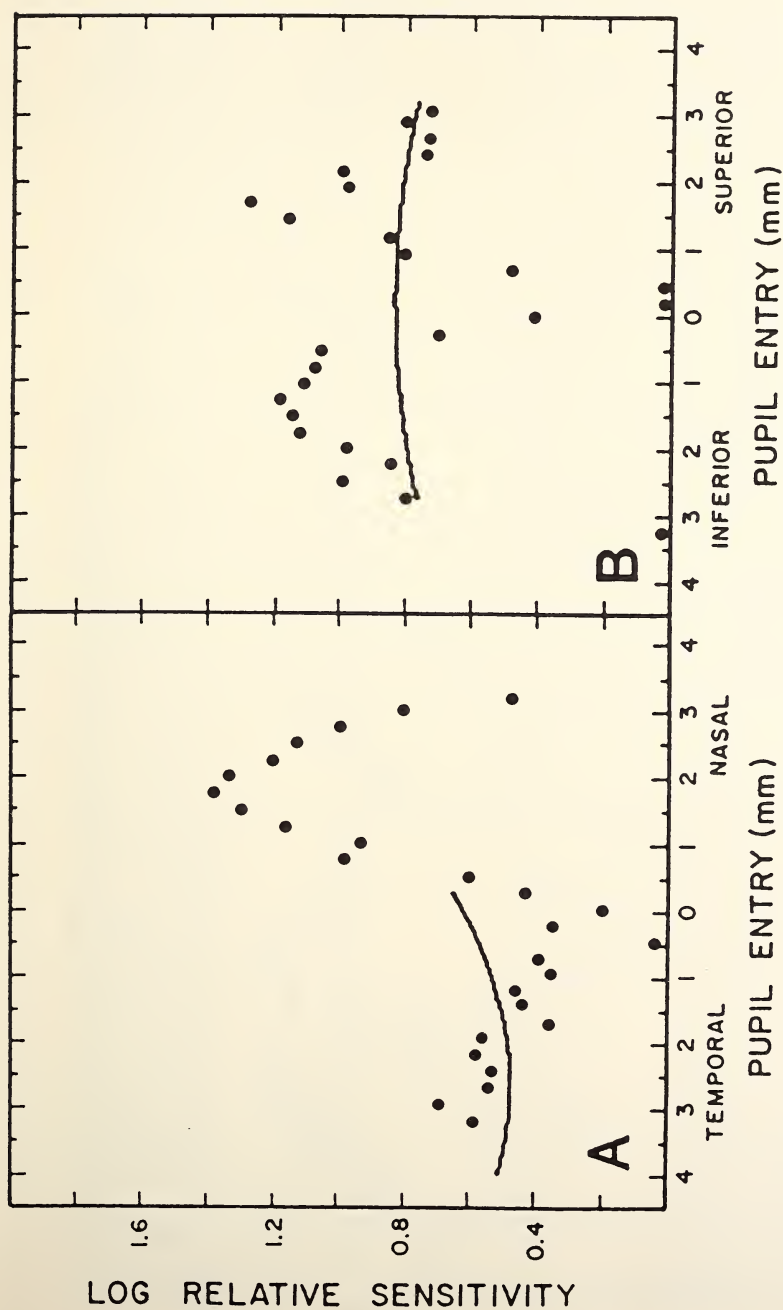
Figure 6. The horizontal (Panel A) and vertical (Panel B) SCE measured in an individual having a small dense posterior subcapsular cataract (solid dots). The smooth curve illustrates how poorly the traditional parabolic analysis represents the cataract-contaminated data sets.
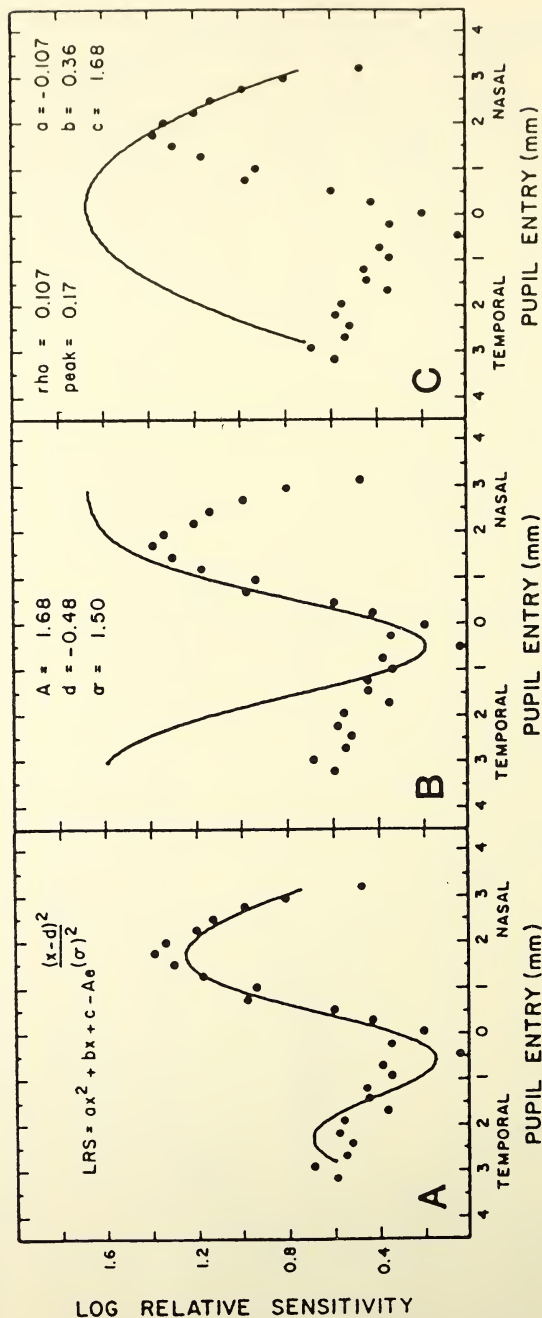
Figure 7. The new two component model of SCE behavior applied to analyze the impact of the small dense posterior subcapsular cataract (illustrated in Figure 6) on the underlying directional properties of the photoreceptors (horizontal meridian). Panel A displays the model's best fitting function (smooth curve) to the cataract-contaminated horizontal SCE data set (solid dots). Panel B illustrates the loss in log sensitivity resulting from the model's best estimate of the cataract maximum density (A), location of maximum density (d) and spread (σ) in the horizontal pupillary meridian. Panel C illustrates the estimated central alignment tendencies (peak location) and distributive properties (rho) of the photoreceptors by showing only the parabolic portion of the model (smooth curve).
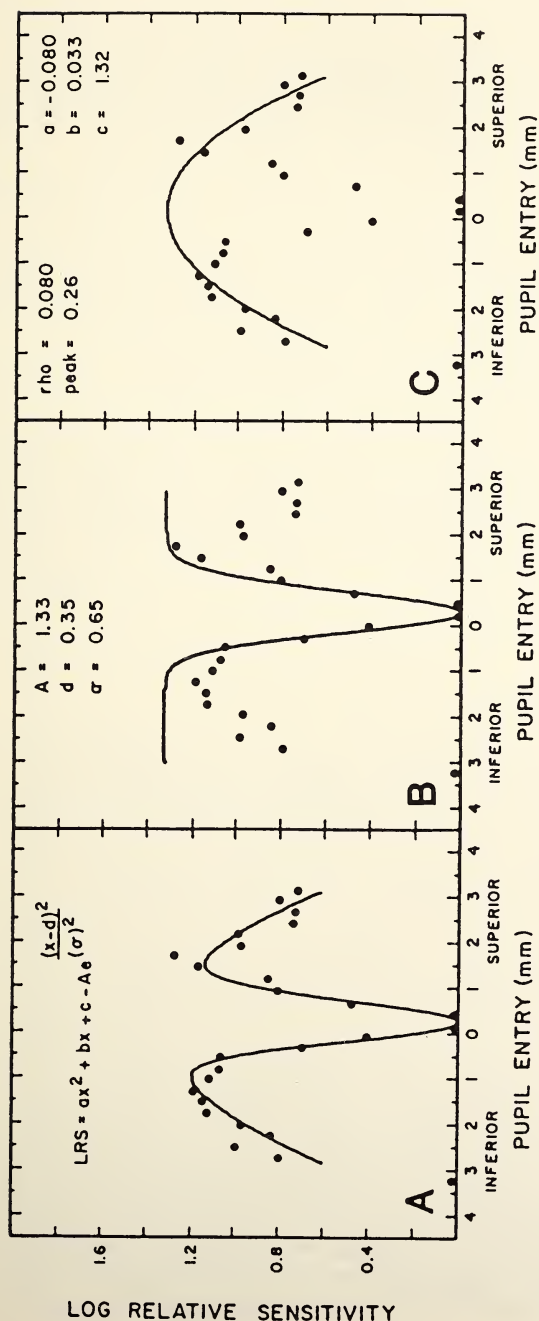
Figure 8. The new two component model of SCE behavior applied to analyze the impact of the small dense posterior subcapsular cataract on the underlying directional properties of the photoreceptors (vertical meridian). Panel A displays the model's best fitting function to the cataract-contaminated vertical SCE data set. Panel B illustrates the loss in log sensitivity resulting from the model's best estimate of the cataract maximum density (A), location of maximum density (d) and and spread (σ) in the vertical pupillary meridian. Panel C illustrates the estimated central alignment tendencies (peak location) and distributive properties (rho) of the photoreceptors by showing only the parabolic portion of the model (smooth curve).

Future Studies

To evaluate whether or not the proposed model of SCE behavior in the presence of cataracts reflects the underlying influence of both the cataract and receptor directional properties, two different experiments will be performed. In the first experiment, measurement will be made of the SCE both before and after cataract removal and intraocular lens (IOL) implantation. Assuming the surgery per se has little or no influence on receptor directional properties, the parabolic portion of the modeled SCE behavior in the presence of cataract will be compared to the traditional parabolic analysis made in the absence of cataract (IOL) to determine how well the model predicts the underlying receptor directional properties.

In the second experiment (suggested by Joel Pokorny), measurement will be made of both photopic and scotopic SCEs in the presence of cataract in the same individual. Since the scotopic SCE is essentially flat, measurement of the scotopic SCE in the presence of cataract should reflect the density characteristics of the cataract as a loss in log sensitivity corresponding to the density, location and spread of the cataract. These characteristics will be compared to the modeled cataract effects obtained during photopic measurements to see how well the model predicts the underlying density characteristics of the cataract.

Limitations and Conclusion

It is important to emphasize that the new model of SCE behavior in the presence of cataracts assumes the underlying directional properties of the photoreceptor reflected in the SCE are well modeled by a parabola and the loss in log sensitivity resulting from the density effects of a cataract are well modeled by a negative Gaussian. The parabolic assumption is well grounded in a long history of experimental observations. There are few data against which to test the assumption of a Gaussian density profile for characterizing cataracts. Further, even if simple cataracts are well modeled using the Gaussian assumption, not all cataracts are simple or well behaved. Some have multiple areas of high density and cannot be well modeled with a simple Gaussian. Even if the negative gaussian is a good model of the loss in log sensitivity resulting from a simple cataract, at this point in our investigations it is not clear what percentage of the vast variety of cataracts can be modeled using the Gaussian assumptions, and what errors will be induced if the Gaussian assumption is applied to irregular cataracts. Nevertheless, it is clear that the effects of cataracts on the measured SCE can be large and well hidden, and while the proposed model may have limitations, it is a significant step toward specifying and solving the problems of interpreting the SCE measured in the presence of cataracts.

Dedication

Both of us have had the privilege of knowing Gordon G. Heath as a scientist, mentor, administrator and friend. It is with gratitude and in friendship that we dedicate this paper to Gordon G. Heath.

Acknowledgements

References

1. Denton, E.J. (1954) On the orientation of molecules in the visual rods of *Salamandra maculosa*. *J.Physiol.* 124:17P-18P.
2. O'Brien, B. (1951) Vision and resolution in the central retina. *J.Opt.Soc.Am.* 41:882-894.
3. Tobey, F.L., Enoch, J.M. and Scandrett, J.H. (1975) Experimentally determined optical properties of goldfish cones and rods. *Invest.Ophthalmol.* 14:7-23.
4. Sidman, R.L. (1957) The structure and concentration of solids in photoreceptor cells studied by refractometry and interference microscopy. *J.Biophys.Biochem.Cytol.* 3:15-30.
5. Barer, R. (1957) Refractometry and interferometry of living cells. *J.Opt.Soc.Am.* 47:545-556.
6. Laties, A.M. and Enoch, J.M. (1971) An analysis of retinal receptors orientation. I. Angular relationship of neighboring photoreceptors. *Invest.Ophthalmol.* 10:69-77.
7. Stiles, W.S. and Crawford, B.H. (1933) The luminous efficiency of rays entering the eye pupil at different points. *Proc.Roy.Soc.Lond.* B112:428-450.
8. Applegate, R.A. Aperture effects on phototropic orientation properties of human photoreceptors, *Ph.D. Thesis* Berkeley: University of California, 1983.
9. Bedell, H.E. Retinal receptor orientation in amblyopic and nonamblyopic eyes assessed at several retinal locations using the psychophysical Stiles-Crawford function, *Ph.D. Thesis* Gainesville: University of Florida, 1978.
10. Enoch, J.M. Retinal Receptor Orientation and Photoreceptor Optics. In: *Vertebrate Photoreceptor Optics*, edited by Enoch, J.M. and Tobey, F.L. Newark, NJ: Springer-Verlag, 1988,

11. Bedell, H.E. and Enoch, J.M. (1979) A study of the Stiles-Crawford function at 35 degrees in the temporal field and the stability of the foveal Stiles-Crawford function over time. *J.Opt.Soc.Am.* 69:435-442.
12. Richards, W. (1969) Saccadic suppression. *J.Opt.Soc.Am.* 59:617-623.
13. Campos, E.C., Bedell, H.E., Enoch, J.M. and Fitzgerald, C.R. (1978) Retinal receptive field-like properties and Stiles-Crawford effect in a patient with traumatic choroidal rupture. *Doc.Ophthalmol.* 45:381-395.
14. Enoch, J.M. and Birch, D.G. (1981) Inferred positive phototropic activity in human photoreceptors. *Phil.Trans.Roy.Soc.Lond.* B291:323-351.
15. Applegate, R.A. and Bonds, A.B. (1981) Induced movement of receptor alignment toward a new pupillary aperture. *Invest.Ophthalmol.Vis.Sci.* 21:869-873.
16. Applegate, R.A., Adams, A.J., Bradley, A. and Eisner, A. (1986) Total occlusion does not disrupt photoreceptor alignment. *Invest.Ophthalmol.Vis.Sci.* 27:441-443.
17. Hamer, R.D., Yasuma, T., Lakshminarayanan, V., Enoch, J.M., Birch, D.G. and Birch, E.E. (1985) Stiles-Crawford functions are not broader after one week of total light exclusion. *Tech.Digest Non-Invasive Assess.Vis.Funct.OSA* WA7-1-WA7-4.
18. Enoch, J.M., Hamer, R.D., Lakshminarayanan, V., Yasuma, T., Birch, D.G. and Yamade, S. (1987) Effect of monocular light exclusion on the Stiles-Crawford function. *Vis.Res.* 27:507-510.
19. Fankhauser, F., Enoch, J.M. and Cibis, P. (1961) Receptor orientation in retinal pathology. *Am.J.Ophthalmol.* 52:767-783.
20. Enoch, J.M., VanLoo, J.A. and Okun, E. (1973) Realignment of photoreceptors disturbed in orientation secondary to retinal detachment. *Invest.Ophthalmol.* 12:849-853.
21. Fitzgerald, C.R., Birch, D.G. and Enoch, J.M. (1980) Functional analysis of vision in patients following retinal detachment repair. *Arch.Ophthalmol.* 98:1237-1244.
22. Smith, V.C., Pokorny, J., Ernest, J.T. and Starr, S.J. (1978) Visual function in acute posterior multifocal placoid pigment epitheliopathy. *Am.J.Ophthalmol.* 85:192-199.
23. Smith, V.C., Pokorny, J. and Diddie, K.R. (1978) Color matching and Stiles-Crawford effect in central serous choriodopathy. *Mod.Prob.Ophthalmol.* 19:284-295.
24. Pokorny, J., Smith, V.C. and Ernest, J.T. (1980) Macular color vision defects: specialized psychophysical testing in acquired and hereditary choroiretinal diseases. *Int.Ophthalmol.Clinics* 20:1,53-81.
25. Fitzgerald, C.R., Enoch, J.M., Campos, E.C. and Bedell, H.E. (1979) Comparison of visual function studies in two cases of senile macular degeneration. *Albrecht Von Graefes Arch.Klin.Exp.Ophthalmol.* 210:79-91.
26. Bedell, H.E., Enoch, J.M. and Fitzgerald, C.R. (1981) A graded disturbance bordering a region of choroidal atrophy. *Arch.Ophthalmol.* 99:1841-1844.
27. Birch, D.G., Sandberg, M.A. and Berson, E.L. (1982) The Stiles-Crawford effect in retinitis pigmentosa. *Invest.Ophthalmol.Vis.Sci.* 22:157-164.
28. Dunnewold, C.J.W. On the Campbell and Stiles-Crawford Effects and Their Clinical Importance, The Institute for Perception, RVO-TNO:Soesterberg, 1964. pp. 1-84.

29. Pokorny, J., Smith, V.C. and Johnston, P.B. (1979) Photoreceptor misalignment accompanying a fibrous scar. *Arch.Ophthalmol.* 97:867-869.

30. Yamade, S., Lakshminarayanan, V. and Enoch, J.M. (1987) Comparison of two fast quantitative methods for evaluating the Stiles-Crawford function. *Am.J.Optom.Physiol.Opt.* 64:621-626.

31. Gass, J.D.M. (1973) Drusen and disciform macular detachment and degeneration. *Arch.Ophthalmol.* 90:206.

32. Chandra, S.R. (1874) Natural history of disciform degeneration of the macula. *Am.J.Ophthalmol.* 78:579.

33. Sarks, S.H. (1980) Drusen and their relationship to senile macular degeneration. *Aust.J.Ophthalmol.* 8:117.

34. Gregor, A., Bird, A.C. and Chisholm, I.H. (1977) Senile disciform macular degeneration in the second eye. *Br.J.Ophthalmol.* 61:141.

35. Smiddy, W.E. and Fine, S.L. (1984) Prognosis of patients with bilateral macular drusen. *Ophthalmol.* 91:271.

36. Sorsby, A. *The Incidence and Causes of Blindness in England and Wales. Reports on Public Health and Medical Subjects,* London: H.M. Stationary Office, 1966.

37. Westat, Inc. *Summary and Critique of Available Data on the Prevalence and Economic and Social Costs of Visual Disorders and Disabilities, Rockville, MD, 1976.*

38. Applegate, R.A. and Massof, R.W. (1985) Interpreting the Stiles-Crawford effect in patients with cataracts. *Tech.Digest Non-Invasive Assess.Vis.Funct.OSA* WA6-1-WA6-4.

NON-INVASIVE 'DISSECTION' OF EARLY STAGES OF EYE DISEASE USING COLOR

Anthony J. Adams
Professor of Optometry and Physiological Optics, and Assistant Dean for Academic Affairs at University of California at Berkeley. He was a member of the National Academy of Sciences/ National Research Council Committee on Vision from 1982 to 1989; and Chair from 1985 to 1988.

ABSTRACT. The use of colored backgrounds and test targets, designed to isolate particular receptor types, has found broad application in both clinical and normal study populations. Such techniques allow selective adaptation of visual pathways (e.g. chromatic vs. achromatic) or selective isolation of one receptor type in the detection of test targets. My colleagues and I have applied the chromatic isolation approach with two different clinical populations (diabetics and patients with glaucoma). These approaches have yielded new information about the vulnerable sites in the retina and visual pathways early in the natural history of each of these diseases.

Diabetes and glaucoma are blinding diseases which together account for a significant percentage of all blindness, particularly after age sixty.

Each can be treated with some success in the early stages and blindness can either be delayed or avoided. However significant loss of vision can occur prior to loss of visual acuity or visual fields.

Our studies reveal a selective loss of sensitivity for light detected by S (blue-sensitive) cones quite early in the natural history of the diseases. At least some pre- and post-receptoral sensitivity loss is involved with diabetes and the presence of macular edema invariably is associated with marked S cone sensitivity losses. In glaucoma the site of large losses of sensitivity in the S cone pathways may be quite different; it is most likely a direct or indirect result of pressure on the axons as they pass through the lamina cribrosa of the optic nervehead. A number of our studies, as well as those of others, have now documented that there can be large S cone sensitivity losses, even in the absence of a visual acuity change, in a variety of retinal vascular, degenerative, as well as inner retina and optic nerve disorders.

NON-INVASIVE 'DISSECTION' OF EARLY STAGES OF EYE DISEASE USING COLOR

Historical perspective

Anomalies or pathologies of the visual system have, historically, provided the researcher a form of non-invasive dissection of the physiology and function of the normal visual system. Conversely, techniques designed to identify the sites of visual processing have been most useful when applied to abnormal pathological conditions; such approaches can provide the clinician with important diagnostic, prognostic, and management information (e.g. monitoring the efficacy of some new therapy).

The use of color and the measurement of color discrimination has been an important approach in both of these aspects. The study of dichromats, individuals whose cone photopigment complement is reduced from three to two, has played a central role in the understanding of normal color vision. In the early 1960s the eyes of these individuals provided the first direct measurements of the kinetics and spectral sensitivity of cone photopigments in the human eye.[1,2] Now approaches in molecular genetics have identified not only the chromosomal location but the nucleotide sequence of each of the genes encoding each of the human photopigment.

Both color discrimination data and data involving use of colored stimuli, where color discrimination is not required or central to the task, have been used in the study of clinical variations of visual perception and visual pathway function. Perhaps the oldest use of color to clearly differentiate receptor function is found in the measurements of dark adaptation: small red targets can be used to isolate cone function, while large blue-green targets, focused onto peripheral retina, are used to identify rod function. Clinicians have long used this to establish normal function of the rod and cone receptors in suspected eye disease, and researchers have used it to ensure testing of a particular class of receptors.

The use of colored backgrounds and test targets, designed to isolate particular receptor types, has found broad application in both clinical and normal study populations.[3] Such techniques allow selective adaptation of certain visual pathways (e.g. chromatic vs. achromatic) or selective isolation of one receptor type in the detection of test targets.

Application to two common disorders: rationale

In this paper I show how my colleagues and I have applied the chromatic isolation approach with two different clinical populations (diabetics and patients with glaucoma). These approaches have yielded new information about the vulnerable sites in the retina and visual pathways early in the natural history of each of these diseases.

Need for early detection

The traditional measures of vision function (visual acuity, visual fields) are most useful in the documentation of a patient's visual performance. They often fail, however, to detect more subtle vision changes which are associated with early stages of eye disease. For many ocular disorders the early detection of the disorder can have profound effects on the management and ultimate outcome of the disease. The two diseases, diabetes and glaucoma (ganglion cell axon death related to raised intraocular pressure) are blinding diseases which together account for a very significant percentage of all blindness, particularly in the age group between forty and sixty. Today, each disease can be treated with considerable success in the early stages and blindness can either be delayed or avoided. Yet in each of these conditions significant loss of vision can occur prior to loss of visual acuity or visual fields.

The treatment of diabetic retinopathy has been aimed at retarding the progress and limiting the severity first of proliferative retinopathy[4] and, more recently, the pre-proliferative phases.[5] Unfortunately, retinopathic changes seen ophthalmoscopically, and capillary damage shown by fluorescein angiography may be relatively late manifestations of the vascular complications which compromise the retina. Because of the devastating effects that diabetic retinopathy can have on vision, the identification of precursors to the breakdown of the blood-retina barrier would be highly desirable. Sadly, standard clinical measures of vision (e.g. Snellen acuity, visual fields) are quite insensitive in detecting vision change even *after* the onset of retinopathy.

In glaucoma, loss of nerve fibers, clinically seen as a change in color or shape of the optic nerve head, produces characteristic visual field defects in the peripheral retina. Clinical testing conventionally involves the measurement of peripheral visual fields, the examination of the optic nerve head by ophthalmoscopy, and the measurement of the intraocular pressure which is thought to be related to the axon death. The recent discovery that glaucoma patientss with only mild or moderate visual field defects may have lost as much as 50% of their ganglion cell axons at the optic nerve head[6] has stimulated interest in identification of the earliest vision changes that precede axon death. Within the last decade or so a number of laboratory

studies have indicated that relatively early in the natural history of glaucoma there are changes of some foveal vision functions including color vision and color increment thresholds (eg Pokorny et al, 1979; Adams et al, 1982)[3,7] and spatial and temporal measures of contrast sensitivity.[8,9] Yet in glaucoma, visual acuity changes typically occur only in the more advanced stages of the disease!

Need for monitoring treatment

For both glaucoma and diabetes, medical and surgical therapies have evolved which halt or retard progression of visual loss. The most recent Early Treatment Diabetic Retinopathy Study (ETDRS) indicates that patients with relatively early stage retinopathy (changes at the back of the eye) and clinically significant edema in the central retina can derive substantial benefit from laser photocoagulation of selected parts of the retina. This focal coagulation treatment reduces the edema and increases the chances of improving visual acuity.[5] The same clinical trial is currently assessing the effects of aspirin and systemic doses of aldose reductase inhibitors on the progression of diabetic retinopathy. Such trials involving patients with very early stage retinal disease, and little or no loss of visual acuity, desperately need additional vision monitors of the efficacy of the treatment.

Rationale for use of colored stimuli

Signals arrive at the ganglion cell layer, via the bipolar cells and lateral retinal interactions at the outer and inner plexiform layers, as a result of signals initiated in rods and/or three different cone types . Each cone receptor contains one of three different photopigments (peak absoption at 430nm, 530nm, and 561nm).[10] Two quite different types of ganglion cell project to the lateral geniculate; both have roughly circular spatially antagonistic receptive fields with center and surround organization. However one group (P alpha cells), representing about 10% of all ganglion cells and an even smaller percentage of foveal ganglion cells, has relatively larger cells bodies, axons and receptive fields, conducts faster with transient responses, has *little or no spectral opponency* (antagonism), high temporal resolution for low spatial frequencies, poor high spatial frequency resolution and projects to the magnocellular layers of the LGN. The other group (P beta), representing the overwhelming majority of ganglion cells and almost the sole representatives of the foveal region of the retina, has smaller cells bodies, axons and receptive fields, is slow conducting with sustained responses, has *strong chromatic opponency*, good high spatial frequency resolution and projects exclusively to the four dorsal parvocellular layers.[11] For purposes of understanding some of the results described for glaucoma, it should be noted that this latter chromatically sensitive group of cells is

of two classes:one, with slightly larger cells and receptive fields[12] and considerably less antagonistic center-surround arrangement, receives antagonistic input from S (blue sensitive) cones and a combination of M and L (red and green sensitive, respectively) cones. The other group receives antagonistic input from M and L cones.

Colored backgrounds and test targets can be combined to isolate the pathways whose signals are initiated in one of three different cone types.[13,14] For example, a bright yellow background selectively reduces the sensitivity of the medium (M) and long (L) wavelength sensitive cones while leaving the S cones unaffected. Under such conditions a deep blue test flash is detected only by the "blue" cones. With sufficiently intense yellow backgrounds and test flashes of sufficiently short wavelength, blue flash intensities of up to 100x test threshold are detected exclusively by S cones. Similarly, intense magenta and blue backgrounds can be used to isolate responses initiated by the M and L cones, respectively. In the normal eye the isolation can be achieved over at least a limited portion of the spectrum for each of the cone types (shaded areas in Figure 1).

It is now clear that many retinal diseases[15] result in selective loss of sensitivity for signals detected by the S cones; little or no loss of sensitivity is seen for the medium and long wavelength sensitive cones when flashes are detected against bright magenta and blue backgrounds in the same subjects.

Results: Diabetes

In studies with diabetics (Figure 1) we have shown that even at the very early stages of this peripheral vascular disease, large and selective sensitivity losses occur in the S cone pathways.[16,17,15,7,18,19]

Post-receptoral involvement

These same losses for short wavelengths are also seen when flashes are presented against a bright white background designed to isolate detection by chromatically sensitive pathways. Under these conditions the chromatic, not the achromatic, pathways are probably involved in detection.[20] Figure 2 shows the spectral sensitivity one finds in the normal eye under these conditions.

When the same sized targets are flickered at 25 Hz and detection is based on detection of flicker, quite different spectral sensitivity is determined and this appears to be due to detection by achromatic pathways.[21] Diabetic eyes, particularly in the early stages of the disease, appear to have little or no loss of sensitivity for signals detected by this pathway (Figure 2). Such data provide an interesting confirmation that the loss of sensitivity seen with slow flashes on either a yellow or white background is not due
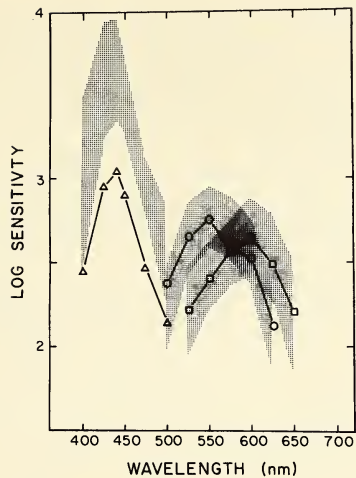
Fig 1. Spectral sensitivities for 34 diabetics under age 40. Open triangles; 4.5 log troland yellow background-S cones. Open circles: 2.9 log troland magenta background-M cones. Open squares 3.2 log troland blue background-L cones. Test spot 1 degree, 1Hz flashes against 10 degree steady background following 3 minutes of adaptation to the background. Shaded area equals 2 standard deviation zone for young age-matched normals with the same 1 and 25 Hz test stimuli.
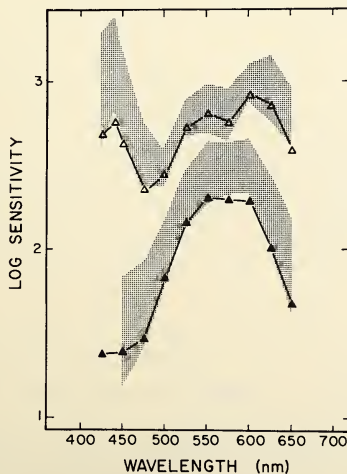
Fig 2. Spectral sensitivity of 14 diabetics (mean age 25 years) with normal (>20/25) visual acuity for a 1 degree spot, flashed against a bright (3 log troland) 10 degree white background. Closed symbols: 1Hz flashes. Open symbols: 25 Hz flashes. Shaded area equals 2 standard deviation zone for young age matched normals with the same 1 and 25 Hz test stimuli.

to the media, since normal sensitivity at short wavelengths for flicker would not be anticipated if the eye contained a pre-retinal yellow filter. In fact the large sensitivity shifts seen in diabetics seem to be due, in part, to changes beyond the receptor (post-receptoral level). More specifically, there is evidence that this loss of sensitivity in the chromatic blue-sensitive pathways , particularly prior to the development of retinopathy, may be a direct result of poor sensitivity control in the outer plexiform layer of the retina, where blue and yellow signals interact in the blue-yellow chromatic pathways.[22,23]

We believe that the post-receptoral sensitivity losses may be among the earliest vision losses we can detect in diabetes, preceding changes in the receptors themselves. What is the basis for this belief? First, S cone photoreceptors are virtually indistinguishable morphologically from M and L receptors, making it difficult to understand, on that basis, why individual S cone receptors themselves would be more vulnerable to insult. That the losses may take place beyond the receptor level is further supported by the fact that some other peculiar characteristics of the S cone pathways (such as their inability to follow fast flicker) are post-receptoral.[24] In addition, the vascular changes in diabetes affect primarily the retinal (as opposed to the choroidal) blood supply, and so very early changes would be expected to take place in the inner retina before the outer retina, i.e. in the post-receptoral layers. We have also shown that the diabetic's S cone pathway sensitivity to a spot of light superimposed on white background (Figure 2) is more nearly normal than when the same spot is superimposed on a yellow background (Figure 1). The Pugh and Mollon[25] model for sensitivity control in the normal S cone pathway suggests that sensitivity control (the attenuation of signals arising from a receptor) can take place at both the S cone itself and at a blue-yellow color opponent post-receptoral site (the "second site") when this second site is polarized or stressed. For example, in this model, after adaptation to a blue or yellow background, sensitivity to a spot of blue light is first lost and then regained via some restoring force (adaptive mechanism). In diabetic patients, this loss of sensitivity does not seem to fully recover, possibly due to some anomaly of the restoring force. We proposed that the relatively normal sensitivity on a white background is due to the fact that the white background, which is neither blue nor yellow, is neutral to the "second site," leaving it unstressed.

I believe we provide further evidence for a post-receptoral component to the loss of sensitivity in two separate experiments.[26-28] The first involves the measurement of the influence of surrounding retina on S cone sensitivity.[26,27] The second involves the measure of unique green hues for which exact cancellation of yellowness and blueness signals are presumed to take place post-receptorally.[28]

According to the Westheimer effect of spatial sensitization,[29] normals show higher sensitivity to a small test probe on a large yellow field than on a smaller one of the same luminance. Since S cones absorb light appreciably only at wavelengths shorter than 550 nm, any changes of S cone sensitivity against an expanded yellow background must be post-receptoral. To assess the post-receptoral function in insulin-dependent diabetics and age-matched normals directly, we measured S cone sensitivity on a large (2°) yellow (588 nm) field and on a small (35' arc) yellow field.[22] If diabetics had normal post-receptoral sensitivity control, we would expect them to show the same amount of spatial sensitization as the controls. We found, however, that three of the diabetics showed a slightly reduced sensitization effect; a fourth showed no effect at all (i.e. sensitivity was the same on both large and small yellow backgrounds). If this reduced sensitization is confirmed in a larger sample it would suggest that for diabetics, as a group, some component of sensitivity loss is due to adaptation at a site beyond the receptor.

The setting of unique green, without blueness or yellowness, is thought to reflect the post-receptoral balance of signals from S and a combination of M and L cones. Consequently a reduced signal from either component receptor output would be expected to shift the wavelength of unique green.

We looked at unique green settings in diabetics with reduced S cone pathway sensitivity.[24] With one exception, the 4 diabetics we've studied had unique green settings that were statistically indistinguishable from 5 normals. The mean wavelength for normals was 501 ± 1.9 nm; for the diabetics it was 504 ± 4.6 nm. The equality of unique green between diabetics, who have S cone sensitivity loss, and normals is consistent with a post-receptoral S cone pathway loss. (We adopted unique green measures, rather than the more common hue cancellation paradigm, because of the former's insensitivity to ocular media filtration.) A selective loss of sensitivity within the S cones themselves would be expected to lead to excess of blue in the hue cancellation paradigm or a shorter wavelength unique green.

Clinical testing formats and results

The measurement of complete spectral sensitivity curves on each of these backgrounds is a time-consuming and difficult task, and is performed in the laboratory in complex Maxwellian view optical systems, generally involving either a bite-bar or chin rest with head restraints. However we have recently reported the use of a relatively simply clinical version of a test which can isolate S cone sensitivity quickly (<5 min/eye) and compare it—in that same time—to the sensitivity of the other cone types.[30] The Berkeley Color Threshold Test (BCT) test is designed to be essentially free
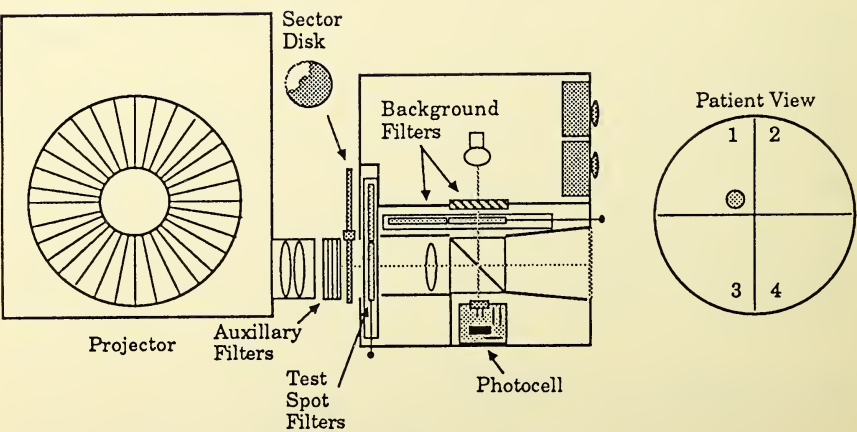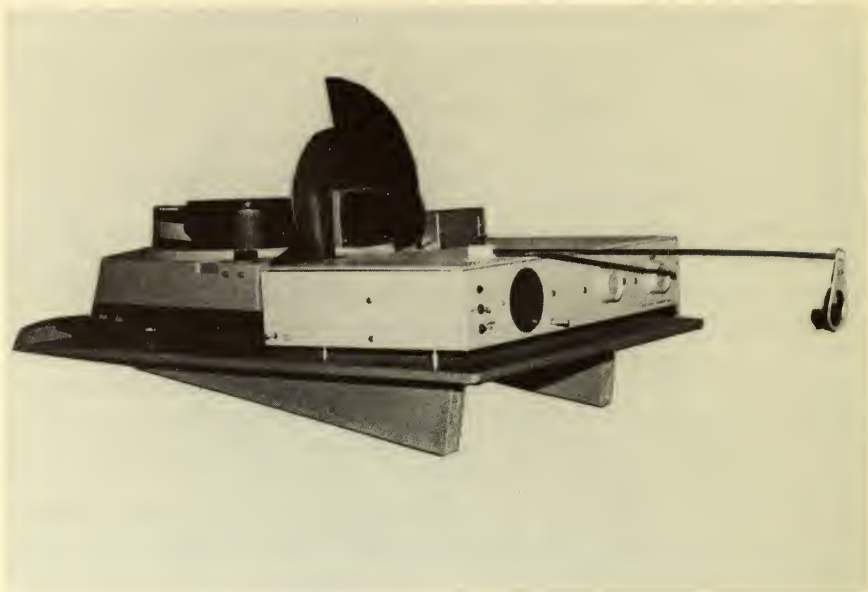
Fig 3. Berkeley color threshold test (BCT) and schematic. In a four alternative forced choice procedure, the patient is presented with a 1 degree spot flashed 'randomly' into one of four quadrants of the background. Each successive presentation is 0.1 log unit less intense until the patient makes one mistake; then the carousel presentation is reversed until the patient again correctly identifies the position of the test flash in the quadrants. Threshold is defined as the intensity of the test flash first correctly identified during the ascending intensity presentation. The viewing screen is rear-illuminated and viewed through an aligning aperture at 33cm.

of differences in transmission through the ocular media and pupil from patient to patient, and from moderate amounts of uncorrected refractive error and fixation unsteadiness (Figure 3).

The test uses a carousel slide projector to provide a four-alternative forced choice presentation of blue or yellows spots against a yellow or blue background (Figure 3). It has already been used in relatively large clinical populations at three different centers to demonstrate that diabetics have a marked loss of S cone sensitivity related to the amount of edema in the retina and to the degree of retinopathy. In one of those studies (Table 1), forty five diabetics, most with 20/20 vision or better, showed selective loss of sensitivity for the S cones. Compared to normals the diabetic group was 40 times less sensitive to blue light on a bright yellow background while being only 2 times less sensitive to yellow light on the same background. Only a 4 times difference could be accounted for on the basis of age differences in the normal and diabetic groups. Even those patients with 20/20 vision had 10 times worse sensitivity than the normal group. Those diabetics with clinically significant edema (retinal thickening within 500µm of the anatomical fovea or having edema within one disk diameter of the fovea and within the 1500µm macular zone) had an average of two lines of Snellen visual acuity loss (1.6 times loss). In comparison they were 63 times less sensitive with their S cones.

Relation of retinopathy to S cone sensitivity loss

In a collaborative study with George Bresnick at the Ophthalmology Dept. at the University of Wisconsin, we used the new test (BCT) of blue cone sensitivity, in a simple clinical format, to evaluate 65 diabetic patients with all degrees of retinopathy as well as 31 age-matched controls.[32] Two other commonly used color vision tests (FM 100-Hue and D-15) were also used for comparison. (Color vision defects on the 100 Hue test are known to be related to the severity of retinopathy.) All of the stereoscopic fundus photographs for the diabetics were graded. Seven retinopathy levels and four macular edema measures (primarily macular edema size) were used in the statistical analysis. The S cone sensitivity measure was predictive of both retinopathy level and macular edema size (p < .001, F-test). Furthermore, only the S cone sensitivity measure discriminated diabetics without diabetic retinopathy from age-matched normal controls. In fact the S cone sensitivity measure was the only significant predictor of the macular edema status.

Our studies following individual diabetics over more than 7 years suggests that the *group* data, which show significant correlation of retinopathy grade to S cone sensitivity, may be deceptive when applied to individuals.

BCT table 1, sensitivity

			Mean Age	Mean years diabetic	Mean VA20/...	Y	B
Diabetics	All Diabetics	n=45	49.8	16.2	25.0	3.49 ± 0.37	2.76 ± 1.02
	No Edema	n=20	43.5	16.1	17.5	3.59 ± 0.28	3.20 ± 0.56
	Edema	n=25	54.8	16.3	31.0	3.42 ± 0.42	2.40 ± 1.16
Diabetics VA ≥ 20/20		n=31	46.1	16.4	16.3	3.62 ± 0.27	3.23 ± 0.47
Normals		n=57	27.0	- - -	>20	3.84 ± 0.16	4.23 ± 0.20

Table 1. Log sensitivity for 45 diabetics with various degrees of retinopathy. Y = yellow test spot ± 1 SD, B = blue test spot log sensitivity ± 1SD. Diabetics with VA 20/20 or better are a subset of the 45 diabetics. Mean visual acuity was derived from log minimum angle of resolution. (From Adams et al, 1987b)

Of the 130 or so patients which have been tested since the study began, 73 enrolled for the extensive increment threshold measures involving approximately 8 hours of vision testing and 3 hours of medical retina examination and photography at each "visit" (actually three separate days). Of those 73 patients, 43 have been involved in multiple followup visits of at least one year separation from the first visit (mean study duration = 3.3 years, range 1 - 7 years). The mean age and duration of diabetes for these patients was 26.9 ± 13.1 and 12.8 ± 7.1 years, respectively; thus most patients were juvenile onset, insulin dependent (IDDM) Type I diabetics. Twenty-five of these patients (mean study duration = 2.5 yrs.) showed no change in retinopathic grade, while 18 patients (mean study duration = 4.1yrs.) did show change in retinopathy level as defined by the categories of diabetic retinopathy described for the Early Treatment Diabetic Retinopathy Study (ETDRS). As might be expected, patients in the study longer were more likely to evidence change. Only two patients developed proliferative retinopathy during the study. Among the remaining patients, for whom there was a retinopathy status change, 6 changed from no retinopathy to mild retinopathy, 10 from mild to moderate, and 5 from moderate to pre-proliferative. Some patients (three) changed through more than one category and they had a mean S cone sensitivity drop of 0.4 log units. However the more striking and more general result is that there is only a small net shift in S cone sensitivity over the mean study duration of 3.3 years for all of these patients. (There was also no significant shift in L or M cone measures.) Excluding the three patients who shifted through more than one retinopathy category, there is no statistically significant difference in net S cone changes between the two groups.

In spite of the lack of correlation between S cone sensitivity measures (or indeed other vision measures which we use) to an individual's shift in retinopathy grade, we have observed that patients do undergo fluctuations in this measure throughout the years of our study. Our impression has been that these fluctuations, which appear to be selective for S cone measures and exceed the variability we find in the same measures for normals, are more related to diabetic control than to fluctuations in retinopathic status. In particular, five diabetics in our study who had large losses in S cone sensitivity (0.6 - 2.9 log units) within 12 hours, either just prior to or just following our vision testing, had either a severe insulin reaction or had to be hospitalized for problems with blood glucose control. On subsequent visits their sensitivities approached their pre-episodic levels. For four of the patients there was no comparable shift in M or L cone measures of sensitivity.

There are a number of implications for a close functional relationship between vision and diabetic control. For example, the neurosensory (vision) deficits may result from metabolic abnormalities in the retina that precede

and contribute to vascular retinopathy. The manifestation of vascular ret-
inopathy, however, may be related to retinopathic signs in a more complex
manner. Studies of the relationship between tight diabetic control and the
progression of retinopathy suggest that in these more advanced cases ret-
inopathy cannot be stopped by continous subcutaneous infusion of insulin
(CSII) provided by the insulin pump. There is some evidence that retinal
function, like some other measures of function (e.g., renal and peripheral
nerve function) is improved with strict diabetic control. We have begun a
collaborative study with John Linfoot on the effects of diabetic control on
vision function. Diabetics who are undergoing intensive insulin therapy
as a result of their recent initial diagnosis or because of poor diabetic control
are tested for S cone sensitivity as they undergo the initial rapid reduction
in blood glucose. The same diabetics are being studied for more sustained
effects of diabetic control (primarily hemoglobin A1c measures) over a
period of three years.

Results: Glaucoma

In glaucoma, a disease often associated with increased intraocular pres-
sure, but always associated with loss of ganglion cell axons leading to
characteristic or "signature" field defects, a different set of results emerges,
suggesting quite different sites for the pathology than in diabetes. Here,
shown in figure 4, sensitivity losses are seen in *both* chromatic and ach-
romatic pathways, though the losses in chromatic pathways appears to be
greater for blue light.[33]

The loss of nerve fibers in glaucoma is thought to involve both a
localized loss at the poles of the optic nerve head, producing the familiar
asymmetric cupping of the disk and notching of the neuroretinal rim, and
loss of axons which produces an overall reduction of sensitivity. This latter
loss may be related to the frequently observed loss of foveal sensitivity
early in the course of the disease and reported by us above.

In a study in collaboration with Roger Husted at the Ophthalmology
Department, Ft. Ord, California, we showed that both ocular hypertensive
and glaucoma patient groups have a generalized reduction in the sensitivity
to short wavelength (blue) targets against a yellow background over the
entire central field region.[34] Sensitivity to yellow light was normal, i.e. not
different from age-matched normal controls (figure 5). We also showed
that a significant number (about 20%) of 32 ocular hypertensives had lo-
calized "glaucomatous-like" field defects with the same blue targets. These
defects were not demonstrable either by conventional perimetry or with
yellow targets on the same yellow bakground. Measurements were per-
formed over the central 30° field. Furthermore, some localized "glaucom-
atous-like" field defects among the 33 glaucoma patients either were not
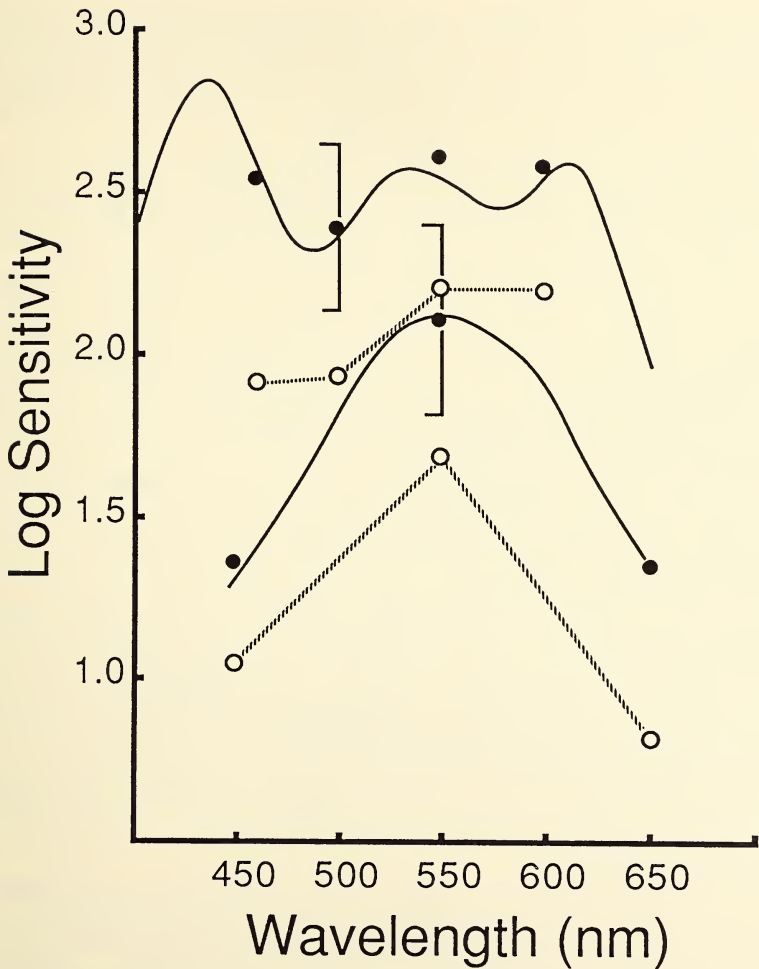
Fig 4. Spectral sensitivity of glaucoma subjects (n = 19, mean age = 51.5 ± 15.8 years) and their age-matched normals (n = 19, mean age = 51.6 ± 15.6 years). Solid lines (through the normal data) are from a separate study of 11 young normals (mean age 24 years) under identical test conditions, except thresholds were determined at 20 nm intervals across the spectrum. Solid symbols and lines: means for age matched normals; Open symbols and dashed lines: means for glaucoma subjects.

Top two functions: Chromatic increment thresholds for a 2° circular target flashed at 1 Hz on a 1270 photopic troland white background, foveally fixated. Bottom two functions: achromatic increment thresholds for a 25 Hz test target under the same conditions. Error bars indicate 1 standard deviation for the normals (SD always less than 0.3 log units at each wavelength).

as extensive, or not revealed, with conventional targets when compared to the blue targets. The loss of sensitivity for both the glaucoma group and the ocular hypertensive group is consistent with Quigley's recent reports,[6] from post-mortem studies of human eyes, of extensive axon death even in the absence of glaucomatous field defects.

In a second phase of this study,[31] we used a series of simple and rapid clinical tests of color and spatial contrast sensitivity to compare the foveal sensitivity of glaucoma and ocular hypertensive patients to age-matched normals. In the first test we used the BCT test of S cone sensitivity and sensitivity to yellow lights on yellow backgrounds to measure the sensitivity of the central macula. For comparison, we administered the D-15, our DSAT version of the D-15 test,[37] the AO HRR color vision book test, and the Vistech version of a spatial contrast sensitivity test.[38] We found the glaucoma group to be about six times less sensitive to blue light than normals, and three times less sensitive than ocular hypertensives (figure 6). These differences were statistically significant and easily identified clinically. There was also a statistically significant difference between normals and ocular hypertensives of about 1.7 times. The results of our Ft. Ord studies prompted a collaborative effort with Chris Johnson and Richard Lewis at U.C. Davis Ophthalmology Department. With them we have initiated a prospective study of the effects of intraocular pressure on measures of blue cone sensitivity in the fovea and peripheral retina for hypertensive and glaucomatous patients. Recently we reported preliminary results showing that, in a group of 38 patients with ocular hypertension and 22 patients with early glaucomatous field defects (all compared to 62 age-matched normals), more abnormalities were obtained for the blue on yellow test condition in 5-10% of ocular hypertension patients and about 15% of the glaucoma patients.[38]

Our studies suggest that a selective loss of sensitivity in pathways fed by S cones (color pathways) and those carrying fast flicker information may be altered before pathways carrying red-green information. We have previously presented evidence from the literature that this is consistent with a sequential loss of large then smaller axons, since the axons carrying red-green information are probably the smallest.[34]

The evidence here, though very preliminary in nature, is that the larger axons of the ganglion cells appear to be preferentially affected by the glaucomatous condition. (Glaucoma is usually accompanied by an increase in intraocular pressure which is thought to produce either direct pressure on the axons as they pass into the optic nerve head and through the lamina cribrosa, or by pressure on the vascular supply to the optic nerve head, thus indirectly affecting the integrity of the gangion cell axons in that location). Thus there appears to be some form of "protection", or sparing, provided to the smaller axons, which carry red-green and high spatial
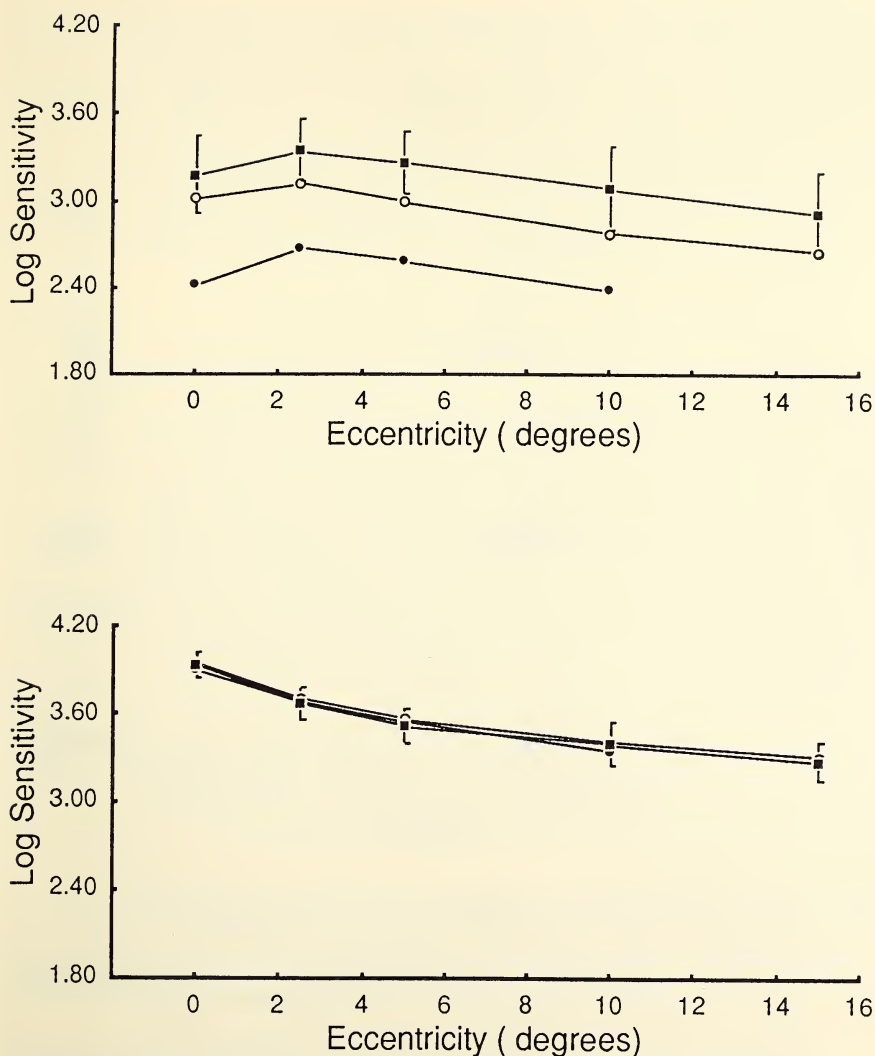
Fig 5. Log sensitivity to 1° blue (top panel) and yellow (lower panel) test flashes (300 ms) on an 11° yellow background at the fovea and at 2.5°, 5°, 10°, and 15° eccentricity. Normal and ocular hypertensive groups measured along the inferior temporal retina, glaucoma group measured along the nasal retina to 10° eccentricity. Sensitivities determined by method of limits. ± One standard deviation is indicated for the normal group. Solid squares = normals, open circles = OHT , closed circles = POAG (primary open angle glaucoma). Mean age and number in group, respectively: normals 59.8 (n = 24), OHT = 61.3 (n = 32), POAG = 65.4 (n = 33).
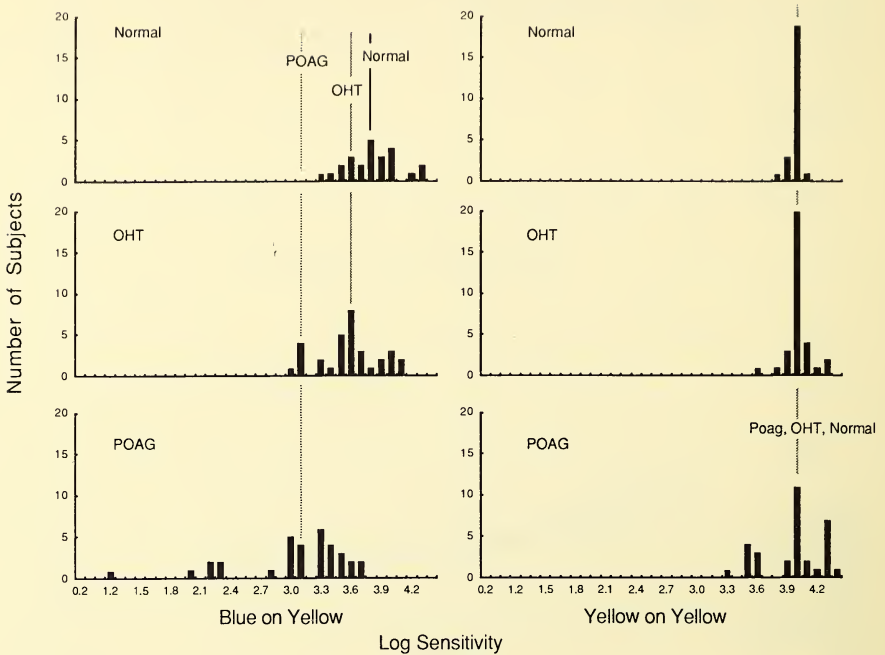
Fig 6.   Frequency histogram for log sensitivities to blue and yellow flashes (1°, 300 ms) on 500 candelas/m² yellow background. Vertical broken lines indicate log mean sensitivity for each group. (OHT - ocular hypertension, POAG - glaucoma.) Test sensitivities are shown for 33 patients with glaucoma, 32 patients with ocular hypertension, and 24 age-matched normals. Each group had the same sensitivity to yellow on a yellow background; there was no statistically significant difference among the groups. The glaucoma group was about six times less sensitive (0.75 log unit) to blue light than the normal group, and three times less sensitive (0.48 log unit) than the ocular hypertension group. These differences are statistically significant (two tailed *t* test, P<.01) and easily identified clinically. There is also a statistically significant 1.7 fold (0.23 log unit) dif-ference between the normal and ocular hypertensive groups (P<.01).

frequency information to the parvocellular layers of the lateral geniculate nucleus. Consistent with this is the fact that visual acuity and contrast sensitivity to high spatial frequencies, along with red-green color discrimination, may be the last functions affected in this disease.

It is interesting to note that there is some convergence of evidence from other forms of vision testing of low spatial and high temporal frequencies, and mesopic increment thresholds, that supports the idea of selective sparing of the smallest axons in glaucoma.

Summary

Our studies of diabetics and glaucoma patients reveal a selective loss of sensitivity for light detected by S cones. The loss in diabetics may be related to sensitivity changes in the receptors themselves but our studies suggest that at least some post-receptoral sensitivity loss is involved. The loss is crudely related to retinopathic status for *groups* of diabetics but not closely related to retinopathic signs, except for macular edema, in individuals. The presence of macular edema invariably is associated with a large S cone sensitivity loss. In glaucoma the use of color also reveals large losses of sensitivity in the S cone pathways but the site of the loss may be quite different, most probably as a result of pressure on the axons as they pass through the lamina cribrosa of the optic nervehead or to the pressure on the vascular tissue in that region. We argue that the concurrent loss of sensitivity for fast flicker, not seen in our diabetic subjects early in their disease, suggests a relative sparing of the smallest axons which carry red-green color information and are responsible for spatial frequency resolution like visual acuity. Changes in these latter functions appear late in the course of the disease.

A number of our studies[40,7,31,39,] as well as those of others[19, 41] have now documented that there can be large S cone sensitivity losses even in the absence of a visual acuity change in a variety of retinal vascular, degenerative, as well as inner retina and optic nerve disorders.

References

1. Weale RA, Vision and fundus reflectometry. Doc Ophthalmol 19:252-286, 1965
2. Rushton WAH, Visual pigments in man, in Dartnall HJA (ed): Handbook of Sensory Physiology, VII/1Berlin, Springer-Verlag, 364-394, 1972
3. Pokorny J, Smith VC, Verriest G and Pinckers AJLG (Editors). Congenital and Acquired Color Vision Defects, Grune and Stratton New York, 1979
4. Rand L I, Prud'homme GJ, Ederer F and Canner PL, Diabetic Retinopathy Study Research Group: Factors influencing the development of visual loss in advanced diabetic retinopathy. Diabetic Retinopathy Study (DRS) Report No. 10. Invest Ophthalmol Vis Sci 6:983-991, 1985

5.   Early Treatment Diabetic Retinopathy Study Research Group, Photocoagu-
      lation for Diabetic Macular Edema, Arch. Ophthalmol. 103: 1796, 1985
6.   Quigley HA, Early detection of glaucomatous damage: II Changes in the
      appearance of the optic disc. Survey of Ophthalmology 30: 111-126, 1985
7.   Adams AJ, Zisman F, Rodic R and Cavender JC, Chromaticity and luminosity
      changes in glaucoma and diabetes, Documenta Ophthalmologica Proceedings
      Series, Colour Vision Deficiencies VI, 33:413-418, 1982a
8.   Arden GB, Jacobson JJ, A simple grating test for contrst sensitivity: Preliminary
      results indicate value in screening for glaucoma. Invest Ophthalmol Vis Sci
      17:23-32, 1978
9.   Tyler CW , Specific deficits in flicker sensitivity in glaucoma and ocular hy-
      pertension. Invest Ophthalmol Vis Sci 20:204-212, 1981
10.   Schnapf JL, Kraft TW, Nunn BJ, and Baylor DA, Spectral sensitivity of primate
      photoreceptors. Visual Neuroscience 1:255-61, 1988
11.   Lennie P, Recent developments in the physiology of color vision. Trends in
      Neuroscience: July 7:243-248, 1984
12.   De Monasterio FM, Asymmetry of on- and off-pathways of the blue-sensitive
      cones of the retina of macaques. Brain Research 166:39-48, 1979
13.   Wald G, The receptors of human color vision. Science 145:1007-1017, 1964
14.   Marre M, Clinical examination of the three color vision mechanisms in acquired
      color vision defects. Mod Prob Opthalmol 11:224-227, 1972
15.   Adams AJ, Chromatic and luminosity processing in retinal disease, Am J
      Optom Physiol Optics 59:642-652, 1982a
16.   Marre E and Marre M, The influence of the three color vision mechanisms on
      the spectral sensitivity of the fovea. Mod Prob Ophthal 11:219-223, 1972
17.   Adams AJ, Zisman F and Cavender JC, Sensitivity loss in chromatic processing
      channels of diabetics, Invest Ophthal Vis Sci, Suppl, 19(4):169, 1980
18.   Zisman F, and Adams AJ, Spectral sensitivity of cone mechanisms in juvenile
      diabetics, Documenta Ophthalmologica Proceedings Series, Colour Vision De-
      ficiencies VI, 33:127-132, 1982
19.   Greenstein VC, Hood DC and Carr RE, A comparison of S cone pathway
      sensitivity loss in patients with diabetes and retinitis pigmentosa. 9th IRGCVD
      Symposium proceedings (in press), 1988
20.   King-Smith PE, Visual detection analysed in terms of luminance and chromatic
      signals. Nature 255:69-70, 1975
21.   King-Smith PE and Cardin D, Luminance and opponent-color contributions
      to visual detection and adaptation and to temporal and spatial integration. J
      Opt Soc Am 66:709-717, 1976
22.   Verdon WA, Adams AJ, Haegerstrom-Portnoy G, Assessment of post-recep-
      toral integrity in diabetics, Invest Ophthal Vis Sci, Suppl, 26(3):308, 1986
23.   Schefrin B, Haegerstrom-Portnoy G, Adams AJ, Verdon W, Hood D Psycho-
      physical testing of sites of disease action secondary to diabetes. Invest Ophthal
      Vis Sci (suppl) 29:68, 1988.
24.   Schnapf J, personal communication, 1988
25.   Pugh EN and Mollon JD, A theory of the pi 1 and pi 3 color mechanisms of
      Stiles. Vision Research 19:293, 1979

26. Haegerstrom-Portnoy G and Adams AJ, Spatial sensitization properties of the blue-sensitive cone pathways, in Color Vision: Physiology and Psychophysics, JD Mollon and LT Sharpe eds, London: Academic Press, 1983, pp. 505-513

27. Haegerstrom-Portnoy G and Adams AJ, Spatial sensitization of the B cone pathways, Vision Research 28(5):,629-638,1988

28. Schefrin BE and Adams AJ, in preparation , 1988

29. Westheimer G, Spatial interaction in the human retina during scotopic vision. J Physiol., Lond. 181:881- 894, 1965

30. Adams AJ, Schefrin BE and Huie K, A new clinical test for eye disease, Am J Optom Physiol Opt 64(1):29-37,1987a

31. Adams AJ, Zisman R, Ai E and Bresnick G, Macular edema reduces B cone sensitivity in diabetes, Appl Optics 26(8):1455-1457, 1987b

32. Witkin SR, Bresnick GH, Friedberg M, Palta M, Adams AJ, Huie KE, Blue cone sensitivity and hue discrimination in diabetic retinopathy, Invest Ophthal Vis Sci, Suppl, 26(3):74, 1986

33. Adams AJ, Rodic R, Husted R and Stamper RL, Spectral sensitivity and color discrimination changes in glaucoma and glaucoma-suspect patients, Invest Ophthal Vis Sci 23(4):517-524, 1982b

34. Heron G, Adams AJ and Husted R, Central visual field measures of blue-sensitive pathways in glaucoma and ocular hypertension, Invest Ophthal & Vis Sci 29(1):64-72, 1988

35. Adams AJ, Heron G and Husted R, Clinical measures of central visual function in glaucoma and ocular hypertension, Arch Ophthalmol 105:782-7, 1987c

36. Adams AJ and Rodic R, Use of desaturated and saturated versions of the D-15 test in glaucoma and glaucoma-suspect patients, Documenta Ophthalmologica Proceedings Series, Colour Vision Deficiencies VI, 33:419-424, 1982

37. Ginsberg AP, A new contrast sensitivity vision test chart. Am J Optom Physiol Opt 61:403-407, 1984

38. Johnson CA, Adams AJ, and Lewis RA, Automated perimetry of short-wavelength-sensitive mechanisms in glaucoma and ocular hypertension; preliminary findings. International Perimetry Symposium, Vancouver Canada, (in press, May 20), 1988

39. Applegate RA, Adams AJ, Cavender JC and Zisman F, Early color vision changes in age-related maculopathy, Appl Optics 26(8):1458-1462, 1987

40. Adams AJ, Selective loss of blue cone mechanism in central serous choroidopathy, Invest Ophthal Vis Sci, Suppl, 22(3):62, 1982b

41. Zwas F, Shin DH, and Mc Kinnon PF, Early diagnosis of glaucoma in ocular hypertensive patients. Invest Ophthalmol Vis Sci 25:193, 1984

DIABETIC DYSCHROMATOPSIA: PREFERENTIAL DAMAGE TO THE BLUE-YELLOW CHANNEL?
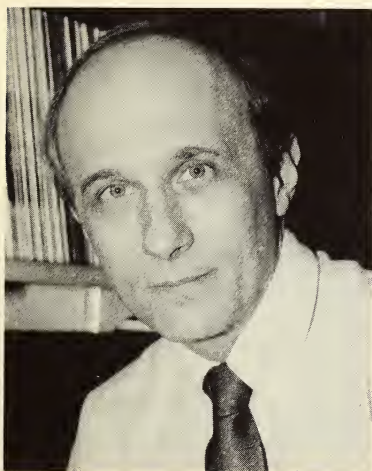
Gary L. Trick
Assistant Professor of Ophthalmology and Visual Science and Assistant Professor of Psychology at Washington University, St. Louis, Missouri. He was chair of the Visual Science Section of the American Academy of Optometry from 1984 to 1986.

ABSTRACT. It is well established that color vision defects frequently occur in diabetic patients. These defects have usually been described as tritan deficiencies indicative of a selective loss of blue-yellow discrimination. This study compared achromatic sensitivities as well as F-M 100-Hue B-Y and R-G partial error scores for a group of diabetic patients and an age-matched control group of normal observers. Results suggest that the color vision deficits in diabetics with little or no retinopathy are not due to selective damage to the B-Y channel but instead to non-selective reduction in both the R-G and B-Y pathways.

DIABETIC DYSCHROMATOPSIA: PREFERENTIAL DAMAGE TO THE BLUE-YELLOW (B-Y) CHANNEL?

Introduction

Of the many visual deficits associated with diabetes mellitus perhaps the most thoroughly studied is the acquired diabetic dyschromatopsia.[1-6] It now is well established that: i) diabetics often exhibit significant color confusion and hue discrimination deficits, ii) these color vision deficits may precede the development of clinically detectable retinal microvascular disease (retinopathy) and iii) these color vision deficiencies generally become more severe as the retinopathy progresses. However, the physiologic basis of the color vision impairments among diabetics, and in particular the dyschromatopsia which precedes the appearance of the characteristic retinal vascular abnormalities, remains unclear.

In order to more fully elucidate the nature of the dyschromatopsia associated with diabetic eye disease, a number investigators have analyzed the diabetic color vision deficits in relationship to the neurosensory mechanisms which mediate normal color vision. Most often these investigations have involved the measurement of either hue discrimination using standardized clinical tests (e.g. Farnsworth-Munsell 100-Hue and D-15 tests) or increment threshold spectral sensitivity.[1-6] From these studies it typically has been concluded that the dyschromatopsia of diabetes is best characterized as a tritan (i.e. blue-yellow) deficit which may reflect preferential damage to a specific pathway involved in the neural processing of color vision information. For example, Bresnick et al.[2] quantitatively examined the red-green and blue-yellow quadrants of 100-Hue results from diabetic patients and detected a predominance of errors in the blue-yellow regions. Similarly, Roy et al.[3] noted that blue-yellow deficiencies were more prevalent than either red-green or mixed defects in early diabetic retinopathy. In increment threshold studies, Adams and co-workers[6] have detected a loss in sensitivity to short-wavelength light among diabetics.

The demonstration of a differential susceptibility of a particular color vision pathway to the vascular and metabolic changes associated with early diabetic retinopathy has implications for the pathogenesis of the visual

dysfunction in this disease. However, it is important to note even in individuals with "normal" color vision blue-yellow discrimination may be poorer than red-green discrimination. This is especially true individuals over 30 years of age. Since it has not been established that the reduction in blue-yellow discrimination among diabetics exceeds the red-green discrimination by an amount in excess of the magnitude of the difference between blue-yellow and red-green discrimination in color normals, the available hue discrimination data must be considered inconclusive evidence for a differential loss in sensitivity. Although the increment threshold results are not necessarily subject to the same criticism, they remain equivocal because reductions in sensitivity to middle and long wavelengths can occur in conjunction with the short wavelength sensitivity decreases.[6] Therefore, this study was designed to quantitatively assess color vision in diabetics and age-matched visual normals using measures of both hue discrimination and increment threshold spectral sensitivity.

Methods and Materials

Hue discrimination: Monocular hue discrimination measurements were obtained with the Farnsworth-Munsell 100-Hue test administered under standard illuminant C (78.5 cd/m^2) with no time limit imposed. The order of presentation of the boxes was varied randomly between patients. An investigator who was not involved in the patient's clinical assessment scored the 100-hue results.

Measurements were obtained from 64 adult subjects with no ocular pathologic disorders or systemic disease (control group) and 75 patits with diabetes mellitus. The mean age of the controls (35.7 ± 10.9 years) was not significantly different from the diabetics (39.2 ± 10.8 years). In the group of diabetic patients 24 participants (16 type I and 8 type II) had evidence of mild to moderate background retinopathy (defined as grade 1a to 1b lesions according to the modified Airlie House classification) determined from seven-field color stereo fundus photos evaluated independent of this investigation. The remaining 51 diabetic patients (28 type I and 23 type II) had no evidence of retinopathy on either ophthalmoscopy or fundus photography. No macular edema was evident in any of the diabetic patients. The average duration of diabetes was 7.8 years for the patients with no retinopathy and 16.9 years for the individuals with background retinopathy. All participants were required to have visual acuity of 20/30 or better and intraocular pressure less than 21 mm Hg.

Two different methods of quantifying the magnitude and type of errors on the test were employed. In the first method the total error score as well as the partial error scores for the blue-yellow (caps 1-12, 34-54 and 76-84) and red-green (caps 13-33 and 55-75) quadrants were assessed as suggested

by Smith, Pokorny and Pass.[7] Since the distribution of F-M 100-Hue error scores for visually normal individuals is non-gaussian the square roots of each participant's total error score, as well as both partial error scores were calculated.[8] The second scoring technique was based upon the method of analysis suggested by Vingrys and King-Smith.[9] This method involves the calculation of a moment of inertia from the color difference vectors which describe each participant's cap arrangement. From this quantitative analysis we examined both the confusion angle (which indicates the type of color vision deficit) and the confusion index or C-index (which indicates the severity or degree of the deficit).

Achromatic and blue-yellow increment threshold sensitivity: A modified Humphrey perimeter was used to measure both achromatic and blue test on yellow background (blue-yellow) increment threshold sensitivity at 74 points within the central 30 degrees of the visual field. For achromatic perimetry the thresholds for detecting a white flash on a 315 apostilb (101.8 cd/m²) white background was determined. For the blue-yellow perimetry a blue, lowpass (< 480 nm) interference filter (Ealing #35-5263) was inserted in the test beam and a yellow Wratten filter (#8) was placed over the background source. With this configuration the yellow background was 216 apostilb (69.4 cd/m²). For each of the 74 points the achromatic thresholds were corrected for age using the sensitivity values included in normative database available with the Humphrey STATPAC. Finally, the difference between achromatic and chromatic thresholds at each test point was computed.

Increment threshold measurements were obtained from 14 adult subjects with no ocular pathologic disorders or systemic disease (control group) and 23 of the diabetic patients who participated in the hue discrimination experiments mellitus. The mean age of the controls (35.7 ± 12.3 years) was not significantly different from the diabetics (44.5 ± 13.4 years). In the group of diabetic patients 6 participants had evidence of mild to moderate background retinopathy (defined as grade 1a to 1b lesions according to the modified Airlie House classification) determined from seven-field color stereo fundus photos evaluated independent of this investigation. The remaining 17 diabetic patients had no evidence of retinopathy on either ophthalmoscopy or fundus photography.

Results

Hue discrimination: Analysis of the F-M 100-Hue total and partial error scores revealed that the square root of the total error was significantly higher in the both groups of diabetic patients relative to the controls (Table 1). This confirms the reduction in hue discrimination expected in these two patient groups. Similarly, the square root of the blue-yellow

TABLE 1. Mean (standard deviation) of the square roots of the F-M 100-hue error scores for each patient group. Values in boldface indicate cases which are significantly different from the normals. NR = no detectable retinopathy; BR = background retinopathy.

ERROR SCORES:	*NORMALS*	*DIABETICS*	
		NR	*BR*
TOTAL	6.71 (2.85)	**9.28 (3.91)**	**10.91 (3.89)**
BLUE-YELLOW	5.13 (2.12)	**7.03 (2.96)**	**8.34 (2.70)**
RED-GREEN	4.24 (2.18)	**5.99 (2.71)**	**6.89 (3.16)**
DIFFERENCE	0.89 (1.28)	**1.03 (1.21)**	**1.45 (1.70)**

partial error score was elevated significantly in the diabetic patients, confirming the reduction in blue-yellow hue discrimination among diabetics. However, a significant increase in the square root of the red-green partial error scores also was detected among both groups of diabetic patients. This suggests that red-green hue discrimination is reduced significantly in diabetic patients.

For the normal controls, as well as for both groups of diabetic patients, the blue-yellow partial error scores were greater than the red-green partial error scores. To assess the relative magnitude of the blue-yellow and red-green hue discrimination deficits among diabetic patients, the difference between the square roots of the partial error scores was calculated. In neither group of diabetics was this difference significantly greater than in the control group.

The Vingrys and King-Smith method of scoring the F-M 100-hue also revealed a significant reduction in hue discrimination (elevation of the C-

MAGNITUDE OF COLOR CONFUSION
(F-M 100-hue)
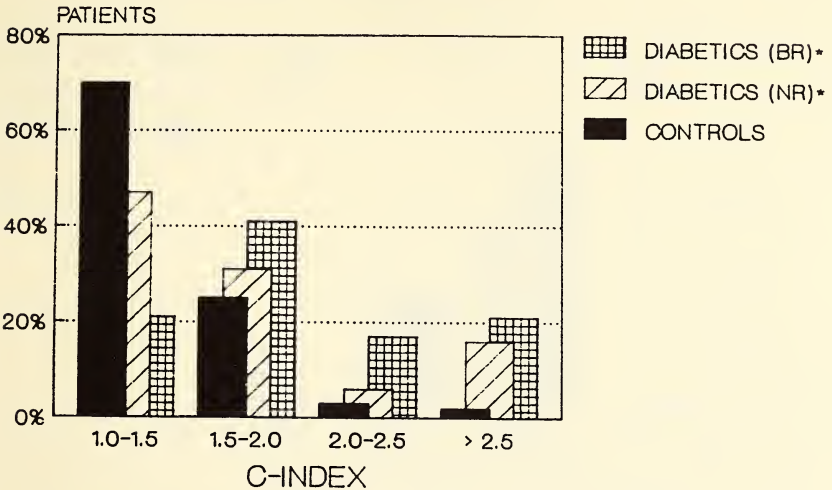
* NR = No Retinopathy
BR = Background Retinopathy

Figure 1. The distribution of the color confusion scores (C-index) for the diabetic patients and the controls is shown. A reduction in color discrimination is evident from the increased C-index of the diabetic patients. Note that significant color discrimination deficits are evident for the diabetics with no detectable retinopathy as well as for the patients with background retinopathy.

index) among both groups of diabetic patients (Fig. 1). However, there was little difference in confusion angle between the controls and either group of diabetic patients (Fig. 2). If only the diabetic patients with significant hue discrimination deficits are considered (i.e. C-index > 2.03, a value that is two standard deviations above the mean C-index [1.39] for the controls), the confusion angles of the diabetic patients fall completely within the range of the controls. Thus, when this method of analysis is employed there is no evidence of a selective blue-yellow deficit in either group of diabetic patients.

Achromatic versus blue-yellow increment thresholds: For both the controls and the diabetic patients the age-corrected achromatic sensitivity was plotted as a function of the difference between achromatic and blue-yellow sensitivity (Figs. 3-5). For the control subjects both the age-corrected achromatic sensitivities and the differences between achromatic and chromatic sensitivities tended to vary randomly between -10 db and +10 db. Among

Gary L. Trick

Figure 2. The distribution of the color confusion angles for the diabetic patient and the controls is shown. The axis of confusion (angle) provides an indication of the type of color vision defect. In the control group the angle generally varied between 40 and 80 degrees. This also was the case for the majority of the diabetic patients including all of the diabetics with color confusion scores more than 2.0 standard deviations above the mean for the control group (data not shown).

the controls no correlation between age corrected sensitivity and the difference between achromatic and chromatic sensitivity was evident when paired points (i.e. test points from the same retinal location) were plotted (Fig. 3). For the diabetic patients both age-corrected achromatic sensitivity and the differences between achromatic and chromatic sensitivity varied over a greater range (Fig. 4), consistent with deficits for both the achromatic and the blue-yellow tests. In addition, among the diabetics the relationship between paired points was not random. Rather there was a significant negative correlation ($r = -0.66$, $p < .001$) between the age-corrected achromatic sensitivities and the difference between achromatic and chromatic sensitivity (Fig. 4). This trend, which suggests that reductions in blue-yellow sensitivity tend to parallel achromatic sensitivity deficits, was evident for individual diabetic patients (Fig. 5) as well as for the grouped data.

```
                                        1

        1                       1
  10 ─┼─    1               2       1   1
   │                  1   1      1 2 8 1 1 1 1 1
   │                    2    12 7 11 10 8 3 2 2 2
   │              1 1    2 8 17 12 19 11 11 8 4 2 2 1
ACHROMATIC        1 2   1 1 1 4 16 16 24 21 15 14 14       3
MINUS     0 ─┼─   1 2   2 2 2 6 8 12 19 37 22 29 10 7 3 2    3
CHROMATIC              2 3 6 7 13 26 40 29 21 16 6 7 2 1      2 2
   │            1 1 2 1 5 13 17 27 31 32 13 10 6 6     3         1 1
   │                1 5 13 12 20 17 18 14 8 4 2 1 1    1
   │          1        1 2 9 6 12 7 4 5 3 2 3 1
  -10 ─┼─              1 4 3 6 4 3 1              1
   │                      1 1
   │                        1
   └── ─ ─ ┼ ─ ─ ─ ─ ─ ─ ┼ ─ ─ ─ ─ ┼─ ─ ─ ─┘
         -10             0            10
        CONTROL MINUS ACHROMATIC
```
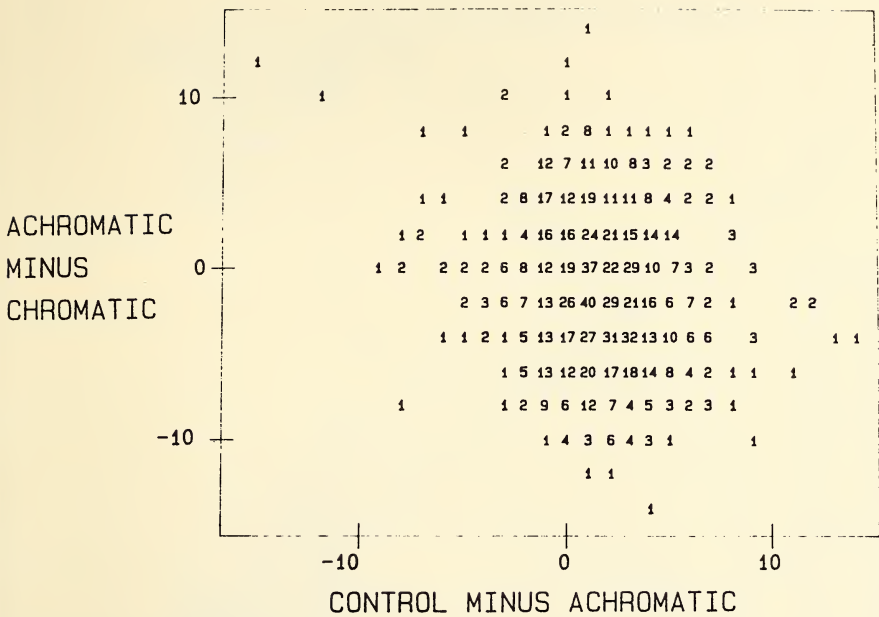
Figure 3. Thresholds for achromatic and chromatic perimetry at each of 74 points in the visual field are shown for the pooled data from 14 control subjects. Each entry represents a comparison of the achromatic and chromatic thresholds for the same retinal location in one individual. The x-axis represents the difference (in db) between the obtained threshold and the average threshold for age-matched observers as derived from the Humphrey STATPAC norms. The y-axis represents the difference (also in db) between the observers achromatic and chromatic thresholds. For these observers the data appear to be randomly distributed along both axes and no correlation is evident.

Discussion

Visual loss resulting from retinal microvascular disease is one of the major complications of diabetes mellitus. Often diabetic visual loss is consider to be synonymous with blindness. It has become quite evident, however, that blindness is a relatively late and often preventable complication of the disease process. Nevertheless, other more subtle and less debilitating visual deficits may accompany the development and progression of diabetic retinal vascular disease.[1-6, 10-13] In studying these early visual manifestations one hopes to learn more about the pathophysiologic basis of visual consequences associated with retinal vascular disease while also developing better methods of detecting those individuals most likely to require medical and/or surgical intervention.

Although it is well-established that significant color vision deficits frequently occur in diabetic patients, and that these deficits often can be

Gary L. Trick
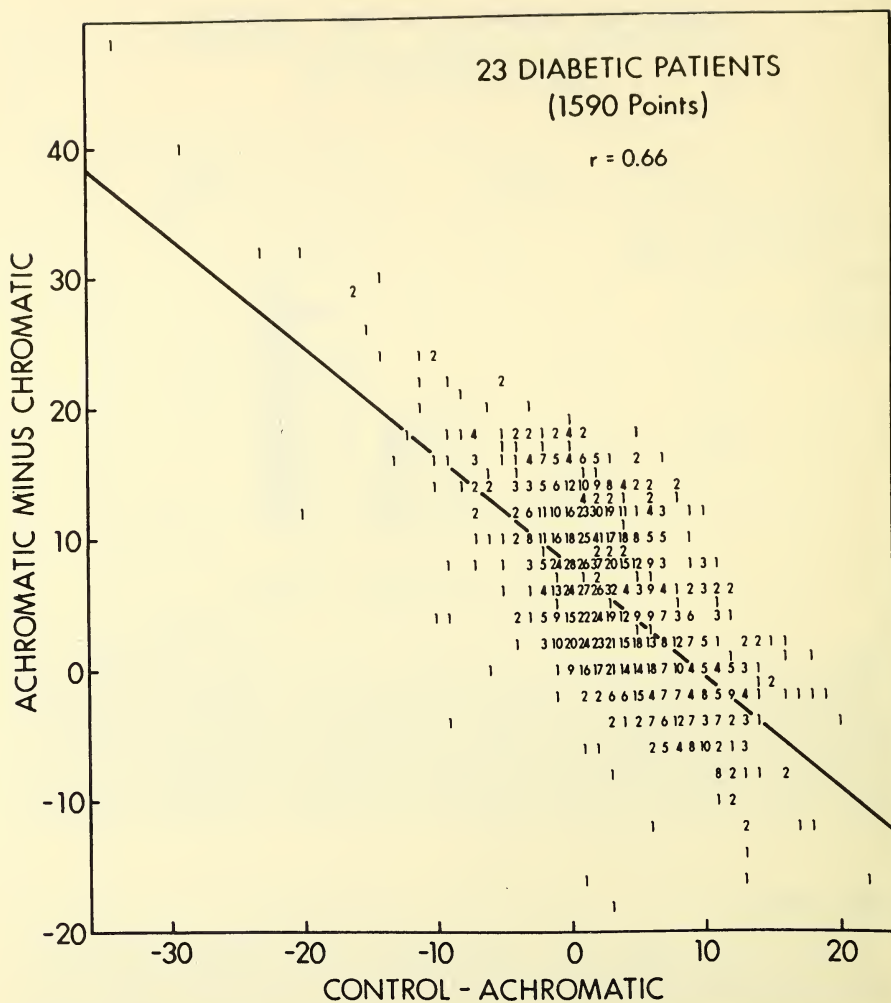
Figure 4. The pooled data for 23 diabetic patients are plotted using the same convention used in Fig 3. In this case, however, a significant negative correlation is apparent and the best fit linear regression (least squares) is shown.

Figure 5A-5C. Identical to Figs. 3 and 4 except the data represent individual diabetic patients.

observed prior to the onset of detectable retinal vascular disease; the nature of the pathophysiologic process mediating the color vision loss remains ill-defined. In this study significant color vision deficits were detected not only in patients with early background retinopathy, but in some individuals these deficits were present before retinal microvascular disease could be detected by either ophthalmoscopy or fundus photography. Therefore, these results support previous suggestions that color vision testing provides a sensitive method for detecting early visual system involvement in diabetes.[1-6]

Previous reports of hue discrimination deficits in diabetics have described tritan deficiencies suggestive of a selective loss of blue-yellow discrimination.[1-5] Unfortunately, in most of these reports the type of hue discrimination deficit has been determined from non-quanitative evaluations of the confusion axes.[1, 3-5] Bresnick et al.,[2] who employed a scoring technique similar to the quadrant analysis used in this investigation, observed that in diabetics the B-Y partial error scores exceeded the R-G partial error scores.

Similar results were obtained in this study. The current results, however, illustrate that i) both the B-Y and R-G partial error scores of diabetics patients are elevated relative to age-matched controls and ii) the difference between the B-Y and R-G partial error scores is similar for diabetics and controls. Therefore, these hue discrimination results are more consistent with the concept of a concomitant reduction in B-Y and R-G discrimination than with the concept of a selective loss of B-Y discrimination. Bresnick et al.[2] may have failed to observe this trend because they did not evaluate their results relative to age-matched controls.

Another interpretation of the hue discrimination results obtained in the present study is that the F-M 100-Hue test is not sensitive enough to reveal selective damage to the blue-yellow channel in patients with little or no retinopathy. This notion is supported by the evidence that blue-yellow deficits have been detected in diabetics with little or no retinopathy when increment threshold techniques are used[6] and by the observation that in previous hue discrimination studies patients with more advance retinopathy were included.[1,4] Therefore, a subset of the patients from this investigation were also tested using blue-yellow and achromatic increment threshold techniques. The results, however, indicate that in diabetics with little or no retinopathy the development of a blue-yellow deficit is generally paralleled by a similar reduction in achromatic sensitivity (Figs. 3-5).

Thus the weight of the evidence from this investigation suggests that the color vision deficits observed in diabetics with little or no retinopathy do not result from selective damage to the B-Y channel. Instead, the early color vision defects in diabetes appear to result from uniform reductions

in both the R-G and B-Y pathways. While significant color vision defects clearly are early signs of visual involvement in the disease process, the mechanisms mediating these visual deficits require further study. Nevertheless, color vision testing remains a viable and useful tool for detecting diabetic eye disease.

REFERENCES

1. Aspinall PA, Kinnear PR, Duncan LJP and Clarke BF., 1983 Prediction of diabetic retinopathy from clinical variables and color vision data. *Diabetes Care* 6:144-148.
2. Bresnick GH, Randall S, Condit RS, Palta M, Korth K, Groo A, Syrjala S. 1985 Association of hue discrimination loss and diabetic retinopathy. *Arch Ophthalmol* 103:1317-1324.
3. Roy MS, Gunkel RD, Podgor MJ., 1986 Color vision defects in early diabetic retinopathy. *Arch Ophthalmol* 104:225-228.
4. Moloney J, Drury MI., 1982 Retinopathy and retinal function in insulin-dependent diabetes mellitus. *Br J Ophthalmol* 66:759-761.
5. Green FD, Gafour IM, Allan D, Barrie T, McClure E, Foulds WS, 1985 Colour vision of diabetic. *Br J Ophthalmol* 69:533-536.
6. Zisman F, Adams AJ., 1982 Spectral sensitivity of cone mechanisms in juvenile diabetics. *Doc Ophthalmol Proc* Ser 33:127-131.
7. Smith VC, Pokorny J, Pass AS., 1985 Color-axis determination on the Farnsworth-Munsell 100-hue test. *Am J Ophthalmol* 100:176-182.
8. Verriest G, Van Laethem J, Uvijls A., 1982 A new assessment of the normal ranges of the Farnsworth-Munsell 100-Hue test. *Am J Ophthalmol* 93:635-642.
9. Vingrys AJ and King-Smith PE, 1988 A quantitative scoring technique for panel tests of color vision. *Invest Ophthalmol* Vis Sci 29:50-63.
10. Della Sala S, Bertoni G, Somazzi L, Stubbe F, Wilkins AJ. 1985 Impaired contrast sensitivity in diabetic patients with and without retinopathy: a new technique for rapid assessment. *Br J Ophthalmol* 69:136-142.
11. Sokol S, Moskowitz A, Skarf B, Evan R, Molitch M, Senior B. 1985 Contrast sensitivity in diabetics with and without background retinopathy. *Arch Ophthalmol* 103:51-54.
12. Arden GB, Hamilton AMP, Wilson-Holt J, Ryan S, Yudkin JS, Kurtz A, 1986 Pattern electro-retinograms become abnormal when background diabetic retinopathy deteriorates to a preproliferative stage: possible use as a screening test. *Br J Ophthalmol* 70:330-335
13. Zingirian M, Polizzi A, Grillo N., 1985 The macular recovery test after photostress in normal and diabetic subjects. *Acta Diabetol Lat* 22:169-172.

A MODEL FOR WAVELENGTH EFFECTS ON BINOCULAR BRIGHTNESS SUMMATION
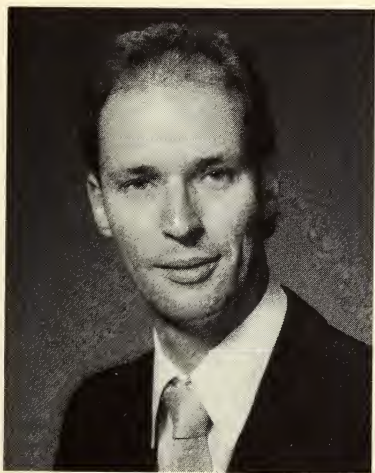
Jeffery K. Hovis
Professor of Ocular Pathology at the School of Optometry, University of Waterloo, Waterloo, Ontario.

ABSTRACT.   The wavelength effects on binocular brightness summation reported by Trick and Guth (1980) are reasonably predicted by a model in which each of the red, green, blue, yellow, and white neural signals has a different nonlinear intensity-response function, and the brightness impressions from each eye are combined by a variation of Engel's (1967) autocorrelation model. This model suggests that wavelength effects on binocular summation may be a result of differences in the achromatic and chromatic intensity-response functions, which occur before the information is combined to give a fused impression.

A MODEL FOR WAVELENGTH EFFECTS ON BINOCULAR BRIGHTNESS SUMMATION

Prior to 1960, the majority of evidence for neural interactions between corresponding retinal areas of each eye was indirect, gathered primarily through psychophysical experiments. However, there is now ample physiological evidence that neural signals from corresponding retinal areas interact at a central level in the visual system (see Reading[1] and Poggio & Poggio[2] for recent reviews). Although the psychophysical studies provide only indirect support for the neural interactions, these experiments do provide an insight into how these interactions relate to perception. One of the earliest hypotheses about the relationship between binocular neural interactions and perception assumes that, if the converging signals from each eye summate, then objects should appear brighter when viewed binocularly. However, the simple experiment of closing one eye when viewing this page and noticing that there is no difference in its brightness suggests that the neural signals do not summate. Nevertheless, in an extensive review of binocular brightness summation literature, Blake and Fox concluded that binocular summation exists for threshold light levels, but, at supra-threshold, binocular summation depends upon luminance and/or contour differences between the stimuli presented to each eye.[3] For example, when equal-luminance lights are presented to corresponding retinal areas, then there is evidence for summation.[4] However, when the luminances of the lights are unequal, as in Fechner's paradox, then the brightness of the fused image is an "average" of the two monocular percepts.[5] This last result implies that the neural interaction can be inhibitory.

In addition to luminance and contour information, two studies have shown that binocular brightness summation also depends upon the wavelengths of the lights presented to corresponding retinal areas.[6,7] Although the retinal illuminances ranged from near photopic threshold in Trick and Guth's[7] experiment to 300 td in de Weert and Levelt's[6] experiment, both studies reported that, as the difference in wavelength of the two lights increases, the degree of binocular summation decreased. That is, as the wavelength difference of the lights increased, the fused image appeared

dimmer. This wavelength effect is similar to brightness summation results reported for monocular viewing.[8]

Another wavelength-dependent result both studies reported is that, the amount of summation varied even when the wavelengths of the lights were identical. For example, a fused image of 425nm lights appeared brighter than a fused image of 575nm lights. However, there is not general agreement on this last result.[9]

A theoretical account of wavelength effects on binocular brightness is difficult because of the wavelength dependency reported for identical lights, and because there is not even general agreement on a model for monocular brightness perception.[7] However, de Weert and Levelt hypothesized that the brightness of the fused image was a weighted average of the two monocular luminances and that, in addition to luminance and contour information, these weighting factors were also wavelength dependent.[5,6]

Recently, the Benzschawel and Guth[10] *ATDN* color vision model has been extended to give reasonable predictions for a number of dichoptic versus monoptic color mixing results.[11] In this paper, the model is developed further to provide a theoretical account of Trick and Guth's binocular summation data. Because of the lack of agreement for a monocular color vision model, this development should only be considered as a first order approximation of the summation data.

Description of the "Binocular *ATDN* Model"

The binocular *ATDN* model is represented schematically in Fig. 1. The first three levels are similar to those of the Benzschawel and Guth model. The cone photoreceptors, L, M, and S, comprise the first stage. They provide input to the second stage mechanisms, which are the nonopponent achromatic mechanism A, the opponent T mechanism (similar to the classical R/G channel) and the opponent D mechanism (similar to the classical B/Y channel). In a departure from all previous opponent colors models, the two outputs of each post-receptor opponent mechanism have different nonlinear responses. That is, the red, green, blue, and yellow hue neural signals each have a different nonlinear intensity response function. This is the critical feature of the monocular *ATDN* model which allows for the monoptic versus dichoptic color mixing predictions. In the present development, a nonlinear response function is also derived for the non-opponent A mechanism to allow for the monocular versus binocular brightness summation predictions. At a central level the achromatic responses from each eye are combined, as are the red-green and blue-yellow responses. The combination rules will be discussed later.

Formulation of the Model

Photoreceptor sensitivities are from Smith and Pokorny as cited in Benzschawel and Guth:[10]

$$L = 0.24\, x' + 0.85\, y' - 0.052\, z', \tag{1a}$$
$$M = -0.40\, x' + 1.2\, y' + 0.084\, z', \tag{1b}$$
$$S = 0.62\, z', \tag{1c}$$

where x', y' and z' are Judd's 1951 modifications of the 1931 CIE color matching functions.

The following equations give the outputs of Benzschawel and Guth's post-receptor mechanisms for the unit radiance spectrum,*

$$A = c_1\, (0.60L + 0.37M), \tag{2a}$$
$$T = c_2\, (0.96L - 1.3M + k_1 S), \tag{2b}$$
$$D = c_3\, (-0.025L + 0.048S). \tag{2c}$$

In eqs 2, A corresponds to the achromatic system, which signals whiteness (or luminance) information, T corresponds to the R/G channel, and D corresponds to the B/Y channel. The k_1 coefficient for S in 2b was adjusted by Benzschawel and Guth to optimize intensity-dependent predictions of small step hue discrimination data, but they suggest that the coefficient should be zero when predicting foveal thresholds. The c_1, c_2 and c_3 coefficients serve as scaling factors and adjust the amounts of A, T, and D that are operated upon by the subsequent nonlinear neural functions.

The nonlinear response functions of Benzschawel and Guth's second stage outputs have the general form,

$$\rho = I^n / (\sigma^n + I^n),$$

where ρ is equal to the hue response, I is equal to the T or D value, σ is the half saturation constant, and n is the exponent. This general form has also been successful in modeling a variety of electrophysiological and psychophysical intensity-response functions for white light (see for example, Hood *et al*[12]).

Their equations for the specific hue responses are:

$$r = T^{0.93} / (T^{0.93} + 1.3) \text{ for } T > 0, \tag{3a}$$
$$g = -(T^{1.1} / (T^{1.1} + 0.83)) \text{ for } T < 0, \tag{3b}$$
$$b = D^{0.75} / (D^{0.75} + 3.0) \text{ for } D > 0, \tag{3c}$$
$$y = -(D^{1.2} / (D^{1.2} + 0.68)) \text{ for } D < 0, \tag{3d}$$

where r, g, b and y represent the red, green, blue, and yellow hue responses. The g and y responses are arbitrarily assigned negative values, whereas the r and b responses are assigned positive values.

---

* To account for some dichoptic color mixing data, the S coefficient in eq 2c was changed to a variable and the M receptor response (with a variable coefficient) was added.[11] However, because the present form of eq 2c has been successful in predicting monocular brightness additivity data near threshold,[13] the original *ATDN* equation for the D response is used in this development.

To predict the brightness summation data, a nonlinear response function for the A mechanism was derived. Although there are a number of possible nonlinear functions, the general form used for the specific hue responses was selected. The nonlinear equation for the achromatic response is:

$$a = A^n/(A^n + \sigma^n). \tag{3e}$$

The exponent, n, and half saturation constant, σ, values were chosen to give the best fit to the data (by visual inspection) using the combination rules defined in the next section.

Rule for Combining Monocular Brightness Signals

To a first order approximation, brightness of the *monocular* impression is defined as the vector sum of the achromatic, red-or-green, and blue-or-yellow responses in both the monocular and current binocular *ATDN* models.[8,10,13] The brightness for a colored stimulus viewed monocularly is expressed as,

$$\psi_M = (a^2 + t^2 + d^2)^{1/2}, \tag{4}$$

where a equals the achromatic response from eq 3e, t equals the red-or-green response from eqs 3a or 3b, and d equals the blue-or-yellow response from eqs 3c or 3d.

Although there are a number of binocular brightness models,[5,14] a variation of Engel's[15] autocorrelation model is used in this initial development as the rule for combining the monocular brightness impressions. According to Engel,[15] the brightness of a fused image is the weighted vector sum of the two monocular brightness impressions. Mathematically, this is expressed as,

$$\psi_B = [W (\psi_R^2 + \psi_L^2)]^{1/2}, \tag{5}$$

where ψ_B is the brightness of the fused impression, ψ_R and ψ_L are the brightness impressions from the right and left eye, and W is a weighting factor that equals the ratio of the monocular squared autocorrelation functions. This ratio depends upon contour and contrast information presented to each eye. Substituting eq 4 for each monocular brightness impression in eq 5 and then rearranging terms, the brightness of a fused colored stimuli is:

$$\psi_B = [W (a_R^2 + a_L^2 + t_R^2 + t_L^2 + d_R^2 + d_L^2)]^{1/2} \tag{6}$$

In the present form, eq 6 will not predict wavelength effects on binocular summation, because squaring the chromatic responses from each eye changes the g and y responses to positive values. In order to retain their negative values, the squared chromatic responses from each eye are mul-

tiplied by a coefficient, s, that equals 1.0 for r and b responses and -1.0 for y and g responses. The expression for the fused image's brightness becomes:

$$\psi_B = [W (a_R{}^2 + a_L{}^2 + |s(t_R{}^2) + s(t_L{}^2)| + | s(d_R{}^2) + s(d_L{}^2) |]^{1/2} \tag{7}$$

The absolute value of combined red-or-green and blue-or-yellow chromatic signals is necessary to maintain a positive value for the sum of the achromatic and chromatic responses.

Predictions

In Trick and Guth's experiment, subjects adjusted the radiance of two individual lights until each matched a very dim comparison annulus surrounding the test fields. This adjustment defined the unit brightness radiances of the individual lights. These radiances were then reduced by 50%, and the two lights were presented either to the same retinal locus in one eye (monocular mixture) or to corresponding retinal areas of each eye (binocular mixture). If brightness summation is complete, then the mixture's brightness would match the standard. However, if brightness summation is incomplete, then the mixture would appear dimmer than the standard, and the subject would have to increase the radiance of the mixture in order to make a brightness match.

To apply the model to the monocular data, the relative radiance of each monochromatic light was calculated using eqs 1-4 so that the vector sum (using eq 4) equaled the small constant of 1×10^{-3}. (Because the maximum response of the nonlinear functions is 1.0, this small constant was selected to represent near threshold responses.) This series of calculations defined the unit brightness radiance of each light. Receptor responses of two equal-brightness lights were then added in order to calculate the achromatic and chromatic responses of the the monocular mixture using eqs 2 and 3. Finally, the a, r-or-g, and b-or-y values were substituted into eq 4, and the relative amount of the equal-brightness mixture was adjusted to produce a vector sum equal to 1×10^{-3}.

For a binocular mixture of the same two lights, the relative amount of the equal-brightness binocular mixture was calculated by substituting the previously determined unit-brightness achromatic and chromatic responses for each light into eq 7, and adjusting the radiances of each light by the same factor so that ψ_B also equaled 1×10^{-3}. The W coefficient in eq 7 was held at constant value of 1. An iterative procedure was used to determine σ and n in eq 3e which gave the best fit to both the monocular and binocular data (by visual inspection) with the scaling constants, c_1, c_2, and c_3 equal to 1.
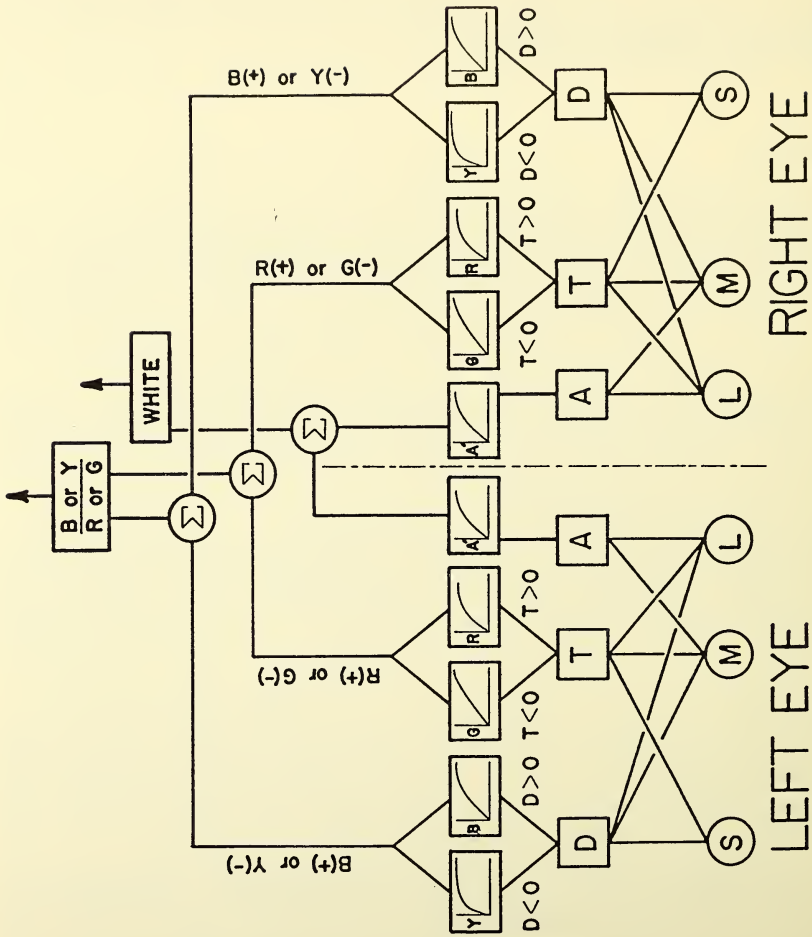
Fig. 1. Schematic of the binocular *ATDN* model described in the text.

RELATIVE AMOUNT WHEN MIXED WITH 650nm
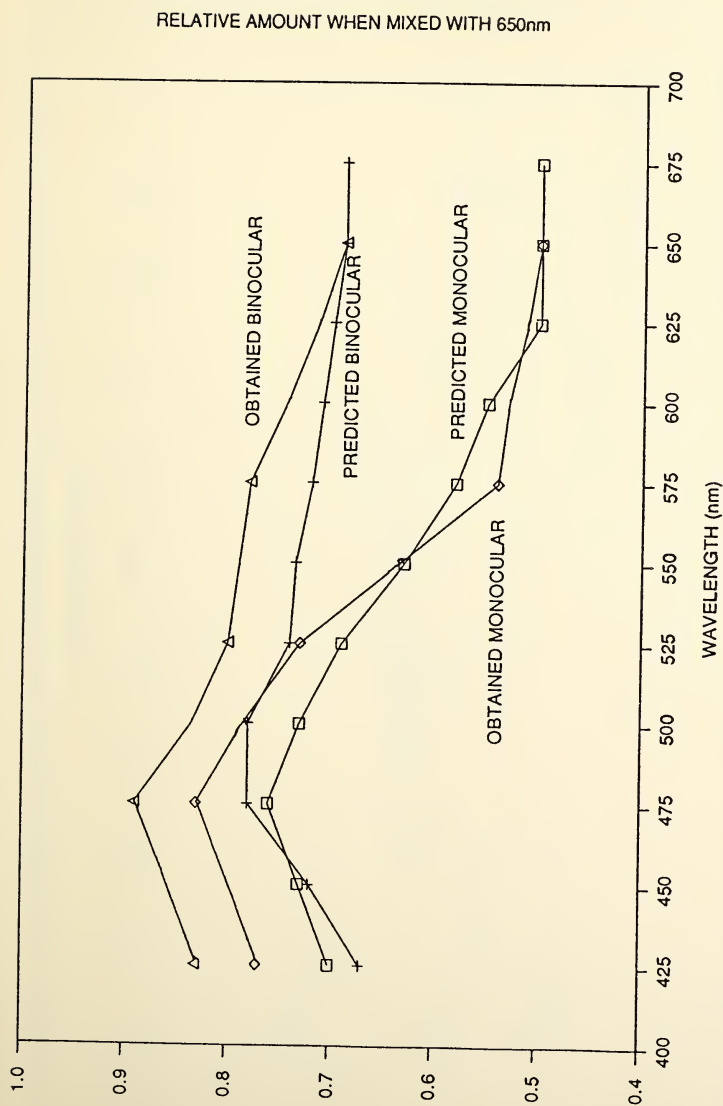
Fig. 2. Obtained and predicted binocular and monocular brightness summation functions for equal-brightness mixtures of the wavelengths listed on the abscissa and 650nm. The ordinate specifies the relative amount of the equal-brightness radiances of the 2 lights in the mixtures required to brightness match a standard. Values at 0.5 indicate complete summation and values greater than 0.5 indicate incomplete summation. Monocular data from Trick and Guth's[7] experiment are represented by the diamonds, and the binocular data are represented by the triangles. Predicted binocular functions are represented by the crosses, and predicted monocular functions are represented by the squares.

RELATIVE AMOUNT WHEN MIXED WITH 425nm
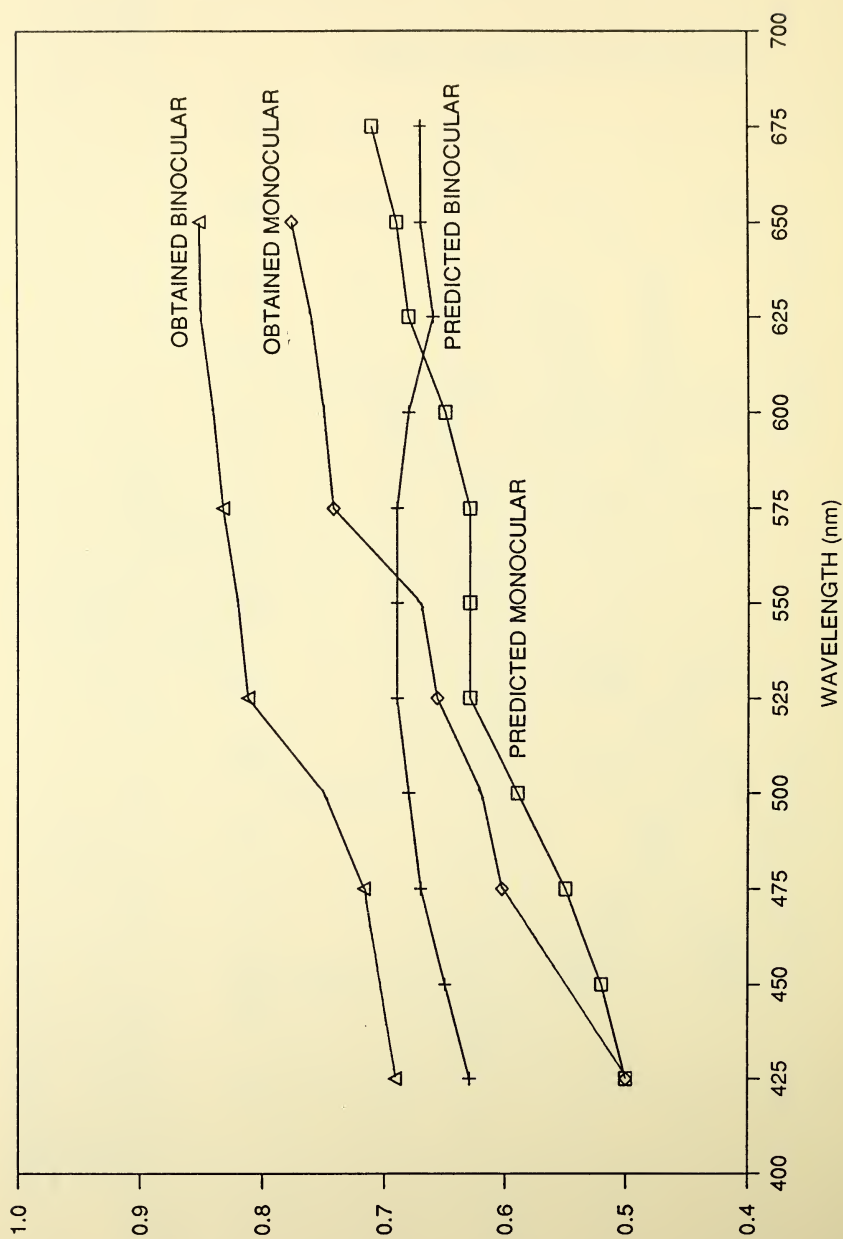
Fig. 3. Same as Fig. 2, except that the wavelengths listed on the abscissa were mixed with 425nm.
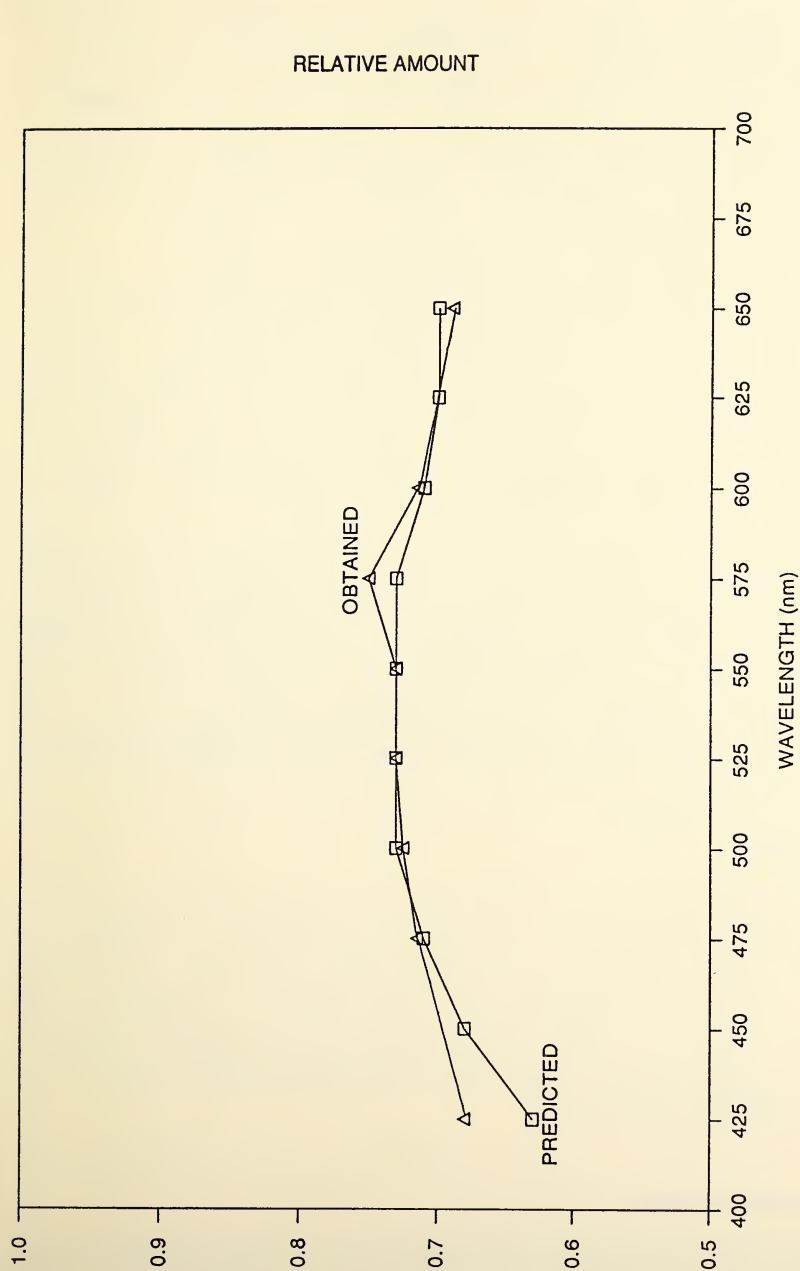
Fig. 4. Binocular summation data (triangles) from Trick and Guth's[7] study when lights with identical wavelengths were presented to corresponding retinal areas. Predictions are represented by the squares. Values at 0.5 indicate complete summation, and values greater than 0.5 indicate incomplete summation.

Data from Trick and Guth's experiment and the predictions are shown in Figs 2 through 4. The σ and n values for eq 3e were 0.52 and 1.1, respectively. Ordinate values in all three figures specify the relative amounts of two unit brightness lights in the mixture required to match the standard's brightness. Points at 0.5 indicate that summation is complete (requiring 0.5 of each unit brightness component) and points above 0.5 indicate incomplete summation. For example, a value of 0.8 for a mixture indicates that 0.8 unit of each light's unit brightness radiance was required to obtain a brightness match when the two lights are mixed together.

Figures 2 and 3 show that, although the model underestimates the amount of monocular and binocular summation for long and short wavelength mixtures, it does provide a reasonable approximation of the data when different wavelengths are combined; however, the model generally overestimates (lower relative amounts as compared with the data). For the red-violet mixtures, the model predicts that, opposite to the obtained results, brightness summation is more complete for a binocular mixture of a red (greater than 610 nm) light and violet (less than 460 nm) light than the monocular mixture.

Figure 4 demonstrates that the model also predicts the wavelength effects on binocular summation when identical lights are presented to corresponding retinal areas in each eye. This last prediction is important, because color vision models with only a linear combination of receptor responses, or with the same exponent and half-saturation constant for the achromatic and chromatic nonlinear intensity-response functions, do not predict these wavelength effects reported for identical lights when using this variation of the vector sum to define the brightness of the fused image.

Conclusion

Although the binocular *ATDN* model can provide reasonable predictions for the wavelength effects on binocular summation, it is possible that other color vision-binocular vision model combinations may also provide equally satisfactory predictions. The important point of this paper is that differences in the achromatic and chromatic nonlinear intensity-response functions, which occur before the information is combined to produce a fused impression, *may* be responsible for the wavelengths effects on binocular brightness summation.

Acknowledgements

References

1. Reading, R.W. 1983 *Binocular Vision: Foundations and Applications.* Butterworths, Boston.
2. Poggio, G.F. and Poggio, T. 1984 The analysis of stereopsis. *Ann. Rev. Neurosci.* 7, 349-412.
3. Blake, R. and Fox, R. 1973 The psychophysical inquiry into binocular summation. *Percept. Psychophys.* 14, 161-85.
4. Fry, G.A. and Bartley, S.H. 1933 The brilliance of an object seen binocularly. *Am. J. Ophthalmol.* 16, 687-93.
5. Levelt, W.J.W. 1968 *On Binocular Rivalry.* The Hague. Mouton.
6. de Weert, Ch.M.M. and Levelt, W.J.W. 1976 Dichoptic brightness combination for unequally coloured lights. *Vision Res.* 16, 1077-86.
7. Trick, G.L. and Guth, S.L. 1980 The effect of wavelength on binocular summation. *Vision Res.* 20, 975-80.
8. Guth, S.L., Donley, N.J., and Marrocco, R.T. 1969 On luminance additivity and related topics. *Vision Res.* 9, 537-75.
9. Abney, W.de W. and Watson, W. 1916 The threshold of vision for different coloured lights. *Phil. Trans. R. Soc. A.* 216, 91-128.
10. Benzschawel, T. and Guth, S.L. 1984 ATDN: Toward a uniform color space. *Color Res. Appl.* 9, 133-41.
11. Hovis, J.K. 1986 Dichoptic Opponent Hue Cancellations. Ph.D. thesis. Department of Physiological Optics, Indiana University.
12. Hood, D.C., Ilves, T., Maurer, E., Wandell, B., Buckingham, E. 1978 Human cone saturation as a function of ambient intensity: A test of models of shifts in the dynamic range. *Vision Res.* 18, 983-93.
13. Guth, S.L., Massof, R.W., and Benzschawel, T. 1980 Vector model for normal and dichromatic color vision. *J. Opt. Soc. Am.* 70, 197-212.
14. de Weert, Ch.M.M. and Levelt, W.J.W. 1974 Binocular brightness combinations: additive and nonadditive aspects. *Percept. Psychophys.* 15, 551-62.
15. Engel, G.R. 1967 The visual processes underlying binocular brightness summation. *Vision Res.* 7, 753-67.

COLOR THEORY AND THE GUTH-INGLING VECTOR MODEL OF COLOR VISION
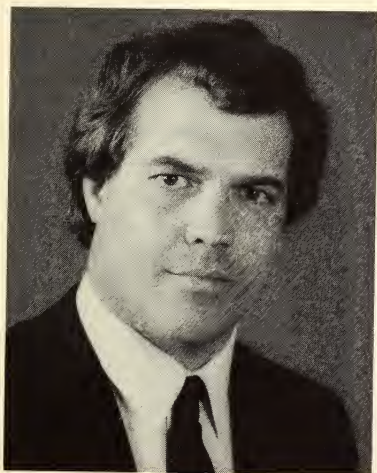
Robert W. Massof
Director of the Vision Research and Rehabilitation Center in the Wilmer Eye Institute at Johns Hopkins Medical Institutions, Baltimore, Maryland. This research was begun while he was a graduate student at Indiana University.

ABSTRACT. The Guth-Ingling vector model of color vision has been remarkably successful in quantitatively simulating absolute and increment threshold spectral sensitivities and heterochromatic threshold additivity failures. This paper reviews the origins and subsequent elaboration of the Guth-Ingling vector model, it provides formal derivations from a general color vision theory and signal detection theory of the two principal operations of the vector model, it discusses and illustrates the formal limitations on applications of the vector model, and it shows how the vector model can be properly extended to wavelength, saturation, and chromaticity discriminations.

COLOR VISION THEORY AND THE GUTH-INGLING VECTOR MODEL OF COLOR VISION

Introduction

Vector models of color vision have become increasingly popular in recent years.[1] Much of this popularity undoubtedly can be attributed to their elegant, but simple designs and to their impressively successful predictions. Best known to the color community, perhaps, are the nearly identical vector models proffered by Guth and his colleagues[2-3] and by Ingling and his colleagues.[4-5] Since the time of these original contributions, other investigators have explored variations and nonlinear elaborations of the basic vector model.[6-8]

Guth's and Ingling's vector models of color vision (hereafter called the Guth-Ingling vector model) consist of two operations. The first operation is a linear transformation of the C.I.E. color mixture primaries (X, Y, Z) to yield three Hering-type color sensory components

$$\begin{bmatrix} A \\ T \\ D \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{1}$$

The second operation is the computation of a vector magnitude from the three-color sensory components

$$d = \sqrt{A^2 + T^2 + D^2}. \tag{2}$$

The Guth and Ingling versions of the model differ only in their choices of values of the matrix coefficients, a_{ij} in Eq. (1), and in *ad hoc* elaborations of the model that will be discussed later.

The Guth-Ingling vector model has been remarkably successful in quantitatively simulating increment threshold spectral sensitivity functions (i.e. t.v.λ. curves) [3-4] and heterochromatic threshold additivity failures.[2-4] The same model has been qualitatively successful in simulating several other sets of color vision data such as color difference thresholds, spectral color-naming functions, spectral saturation functions, brightness-to-luminance ratios, etc.

The purpose of the present article is to analyze the theoretical underpinnings of the Guth-Ingling vector model in order to understand its successful quantitative predictions, to understand the limitations on its performance and applications, and to provide guidelines for further development. In order to accomplish these goals, it will be necessary to apply Bayesian signal detection theory to color vision. Consequently, on a large scale, this theoretical analysis of the Guth-Ingling vector model becomes the vehicle for extending the presentation of a general theory of color vision, the first part of which was described in an earlier paper.[9]

Guth's early development of the vector model was largely theory-free. Later work by Guth and by Ingling focused on extensions of the vector model's predictions and on theoretical interpretations of the vector model parameters. In this paper I will demonstrate that some of these later extensions and elaborations are theoretically sound, whereas many others are not. However, because Guth's original development of the vector model gives it a firm empirical foundation, it is important to first concentrate on retracing his steps.

Origins of the Guth-Ingling Vector Model

Just as color mixture spaces are formal representations of a set of color mixture data, Guth's vector model is a formal representation of his heterochromatic threshold additivity failure data.[10-11] That is, in the original vector model the values of A, T, and D as a function of wavelength were determined by the results of Guth's heterochromatic threshold additivity experiments. Therefore, to fully understand the success of the vector model, it is necessay to understand both the heterochromatic threshold additivity experiment and the representation of the resulting data set by a vector space.

The heterochromatic threshold additivity experiment consisted of measures of a fovel absolute thresholds for mixtures of monochromatic stimuli. This experiment tested the additivity hypothesis (i.e., Abney's law). Explicitly, if e_i is the threshold energy for stimulus of wavelength λ_i and e_j is the threshold energy for a stimulus of wavelength λ_j, then Abney's law would predict that the threshold energy for the heterochromatic mixture, $\lambda_i + \lambda_j$ would be $(e_i + e_j)/2$. Guth's study showed that Abney's law failed for foveal absolute thresholds.

Inspired by the earlier work of Swets[12] and of Cohen and Gibson[13], Guth suggested that monochromatic threshold stimuli might add vectorially. Different wavelength stimuli would be represented by different vectors. Vectors of equal magnitude were defined to represent equally detectable stimuli. The threshold additivity data determined the angle between each pair of stimulus vectors corresponding to each wavelength pair used in the

threshold additivity experiment. If the stimuli were perfectly additive, the resulting angle would be 0°, if the stimuli were perfectly subtractive, the resulting angle would be 180°. Intermediate level of additivity failure would result in angles between 0° and 180°.

To compute the angle between each pair of vectors, Guth assumed that vector magnitude was proportional to stimulus energy and a unit magnitude vector corresponded to a threshold-level stimulus, irrespective of wavelength. The length of a vector for an arbitrary stiumulus, L_i, was expressed as the ratio of the stimulus energy, E_i, to the threshold energy, e_i, of that stimulus, i.e.

$$L_i = \frac{E_i}{e_i}. \tag{3}$$

Guth called the ratio "threshold luminance." In the additivity experiment, one stimulus of wavelength λ_i was set to a threshold luminance of 0.5 (i.e. one-half threshold energy), and another stimulus of wavelength λ_j was added to the half-threshold level λ_i until the mixture came to threshold. The quantity of interest was the threshold luminance of the added stimulus, L_j, required to bring the heterochromantic mixture to threshold. A value of L_j other than 0.5 indicates additivity failure.

The cosine of an angle, θ_{ij}, between two vectors representing the stimuli λ_i and λ_j is

$$\cos\Theta_{ij} = \frac{L_{ij}^2 - L_i^2 - L_j^2}{2L_i L_j} \tag{4}$$

where L_{ij} is the magnitude of the resultant vector for the mixture, L_i is the magnitude of the vector for λ_i, and L_j is the magnitude of the vector for λ_j. For Guth's threshold additivity experiment, $L_{ij} = 1$ (i.e. mixture was brought to threshold), $L_i = 0.5$, and L_j was the dependent variable.*

Guth measured threshold additivity for stimulus pairs of 10 different wavelengths that spanned the visible spectrum. The resulting 10x10 matrix was converted to cosines using equation (4) and symmetric cells were averaged. A 10-variable principal-components factor analysis was performed on the resulting trangular matrix. The factor analysis determined that three orthogonal dimensions were sufficient to account for 93% of the variance in the data.[11]

The factor loadings on the resulting three-dimensions, labeled *A, B,* and *C,* were a function of wavelength and represented the values required to generate a unit magnitude vector. Dividing the factor loadings by the threshold energy at each wavelength, e_i, produced spectral weightings that

---

* In actual practice, Guth set L_i to a nominal half-threshold value, but determined the actual L_i from homochromatic additivity measures, i.e. $\lambda_i = \lambda_i$.* However, for the purposes of discussion we will assume that L_i was always successfully set to 0.5.
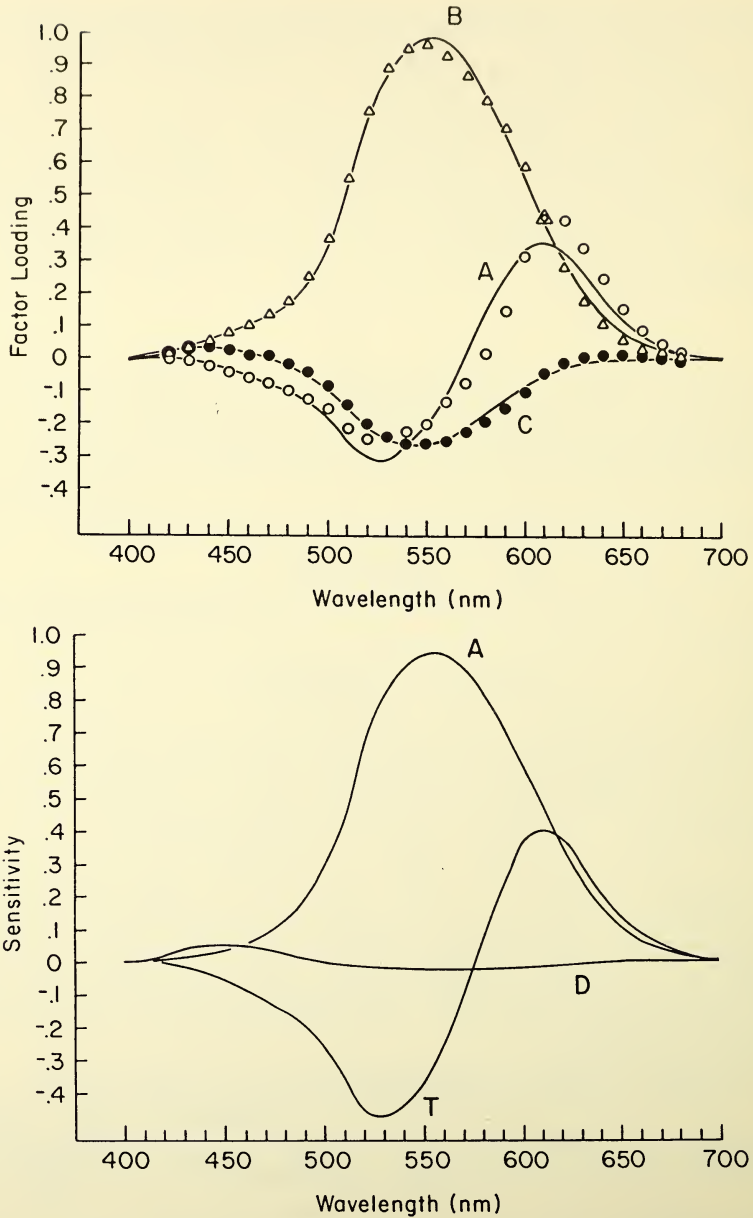
Figure 1. a) Equal Energy *A, B, C* factor loadings as a function of wavelength (symbols), re-plotted from Guth (1969). The solid curves plotted along with the factor loadings are linear transformations of the C.I.E. distribution functions. b) spectral sensitivities of *A, T, D* model vector components, using transformations provided by Guth, Massof, and Benzschawel (1980).

could be thought of as "spectral sensitivities" for the *A*, *B*, and *C* dimensions. Consequently, vector magnitudes, as a function of wavelength for an equal energy spectrum, reproduced the foveal absolute threshold spectral sensitivity data.

The vector space derived from the factor analysis embodied the following properties:
1) Vector magnitude correspondened to stimulus detectability and was proportional to stimulus energy;
2) Vector direction as a function of wavelength was determined by heterochromatic threshold additivity failures (re. equation 4);
3) Spectral loadings on each of the dimensions were determined by the foveal absolute threshold spectral sensitivity function.

The *A*, *B*, and *C* dimensions were only one of an infinite number of sets of orthogonal axes through the same origin that could be used to describe the derived vector space. Figure 1a illustrates the equal energy spectral loadings on the original *A*, *B*, *C* dimensions. Guth noticed that the spectral loadings on dimension *B* were similar to the CIE luminosity function, the loadings on dimension *A* were similar to the action spectra of red-green opponent cells in the monkey lateral geniculate nucleus, and the loadings on dimension *C* were similar to blue-yellow opponent cell action spectra. From Grassmann's color mixture laws, Guth argued that metamers must generate the same vector in the derived space. Since the *A*, *B*, *C* equal energy spectral loadings could be thought of as a set of tristimulus values, Guth reasoned that the *A*, *B*, *C* values must be a linear transformation of the color mixture primaries *(X, Y, Z)*. Within the limits of the data, this relationship was found to be true (re. Figure 1a).

Up to this point the vector space was strictly a formal representation of the data set. By demonstrating that the vector space was a linear transformation of the color mixture primaries, Guth provided the foundation for building his vector model of color vision.[14] In the vector model, Guth defined one dimension to have spectral weightings equal to the C.I.E. photopic spectral luminosity function (i.e. $\bar{y}\lambda$); this dimension was labled *A* for "achromatic". The other two dimensions were linear transformations of the color mixture primaries, designed to have spectral weightings corresponding to a red-green opponent channel, labeled *T* for "tritanopic", and weightings corresponding to a blue-yellow opponent channel, labeled *D* for "deuteranopic". These transformations were constrained to have *T* cross from negative to positive values at 575nm (unique yellow) and to have *D* cross from positive to negative values at 502nm (unique green). Also, the transformations were constrained to provide good predictions of the original additivity failure data. The resulting spectral loadings of the vector model are illustrated in Figure 1b are defined by the linear transformation of the color mixture primaries

$$\begin{bmatrix} A_\lambda \\ T_\lambda \\ D_\lambda \end{bmatrix} = \begin{bmatrix} 0 & 1.0 & 0 \\ 0.7401 & -0.6801 & -0.1567 \\ -0.0061 & -0.0212 & 0.0314 \end{bmatrix} \begin{bmatrix} \bar{x}_\lambda \\ \bar{y}_\lambda \\ \bar{z}_\lambda \end{bmatrix} \qquad (5)$$

Guth's vector model may be regarded as a smoothed and rotated version of his original heterochromatic threshold additivity failure data space, viz. *A, B, C*. Thus, the features built into the data space, threshold additivity failures and threshold spectral sensitivity, are preserved in the highly-constrained vector model.

Guth recognized that the *A, T, D* terms could be expressed as a linear transformation of the foveal receptor primaries, and Massof[15], Ingling and Tsou[4], and Guth, Massof, and Benzschawel[3] described models that explicitly incorporated receptor terms. That is, the foveal receptor primaries (*R, G, B*) are a linear transformation of the color-mixture primaries (*X, Y, Z*).[16] Since the *A, T, D* values are a linear transformation of the color-mixture primaries, *A, T, D* must be a linear transformation of *R, G, B*. This recognition led to physiologically plausible interpretations of the vector model using wiring diagrams such as the one illustrated in Figure 2.

Ingling and Tsou[4] and Guth, Massof, and Benzschawel[3] extended the predictions of the vector model to a variety of t.v. λ. functions on adapting backgrounds of different luminances and spectral compositions. These predictions were accomplished simply by altering the values of the matrix coefficients in Eq. (1). These theorists also attempted to account for color appearance, wavelength discrimination, and saturation discrimination using the vector model. These latter accounts required the invention of *ad hoc* operations, in addition to the operations represented by equations (1) and (2); such ad hoc operations will be discussed later in this paper. First, we will turn our attention to the theoretical analyses of the two operations of the Guth-Ingling vector model, Eqs. (1) and (2), that lead to such successful quantitative predictions.

Linear Transformation of the Color Mixture Primaries

We begin with the general theory of color vision detailed in earlier papers.[9,17-18] To briefly review, the general theory is built on four principal assumptions: 1) visual sensation componency (e.g. Hering opponent colors), 2) quantum absorption componency (e.g. *R, G, B* cones), 3) unique mapping of physical space-time onto sensory space-time, and 4) psycho-neural congruency (i.e. visual sensations arise from physiological operations on light absorptions). These four assumptions are made explicit in the general equation

$$\mathbf{V}(t',x',y') = \Phi\left[\mathbf{Q}(t,x,y)\right] \qquad (6)$$

where V is the visual sensation vector, with components that are functions of sensory time and space coordinates (t', x', y'), Q is the quantum absorption vector, with components that are functions of physical time and space coordinates (t, x, y), and ϕ is the vector of physiological space-time operators that transform the light absorption vector field into the visual sensation vector field.*

For discussion purposes, consider a general visual discrimination experiment often employed in color vision research: a steady and uniform reference stimulus of arbitrary intensity and chromaticity is presented to a specified retinal locus; a difference threshold is measured by asking the observer to detect the difference between the reference stimulus and a steady and uniform test stimulus of variable intensity and chromaticity (in the case of an increment threshold, the test stimulus might represent a physical change in some part of the reference stimulus). Examples of such experiments include absolute, increment, and decrement threshold measures, determinations of Stiles-type threshold-versus-radiance and threshold-versus-wavelength measures, determinations of color difference threshold, absolute and increment threshold heterochromatic additivity studies, etc. (This general visual discrimination experiment may be considered a subset of Brindley's[19] class A experiments.) Due to the constraint of *steady* and *uniform* stimuli, Eq. (6) reduces to the space-time constant functional form

$$V = \phi(Q) \qquad (7)$$

where ϕ is the vector of nonlinear functions that describes the transformation of Q into V for the given space-time parameters. Using Guth's A, T, D notation for the sensory components, in expanded form Eq. (7) becomes

$$V_A = \phi_A(Q_R, Q_G, Q_B) \qquad (8a)$$
$$V_T = \phi_T(Q_R, Q_G, Q_B) \qquad (8b)$$
$$V_D = \phi_D(Q_R, Q_G, Q_B) \qquad (8c)$$

The reference stimulus produces the quantum absorption vector Q_0, which by way of Eq. (7) yields the visual sensation vector V_0. The test stimulus produces the quantum absorption vector Q_1, which differs from Q_0 by ΔQ. Again, by way of Eq. (7) the test stimulus produces the visual sensation vector V_1. The observer is instructed to determine whether V_1 and V_0 are the same or different. Thus, the observer is responding to the difference between V_1 and V_0

---

* See Massof[9] for an in-depth presentation of this general vision theory and the ensuing derivations.

$$\Delta V = V_1 - V_0 = \phi(Q_1) - \phi(Q_0). \tag{9a}$$

Substituting $Q_0 + \Delta Q$ for Q_1 in Eq. (9a), we obtain the difference equation

$$\Delta V = \phi(Q_0 + \Delta Q) - \phi(Q_0). \tag{9b}$$

From the Taylor series expansion

$$\phi(Q_0 + \Delta Q) = \phi(Q_0) + \frac{\partial \phi(Q_0)}{\partial Q} \Delta Q + \frac{\partial^2 \phi(Q_0)}{\partial Q^2} \frac{\Delta Q^2}{2!} + \cdots \tag{10}$$

and considering only small differences. Eq. (9b) can be replaced by the first order approximation

$$\Delta V \approx \frac{\partial \phi(Q_0)}{\partial Q} \Delta Q. \tag{11a}$$

Being reminded that we are partially differentiating a vector with respect to a vector, Eq. (11a) can be rewritten as the system of linear equations

$$\begin{bmatrix} \Delta V_A' \\ \Delta V_T \\ \Delta V_D \end{bmatrix} = \begin{bmatrix} \phi'_{AR} & \phi'_{AG} & \phi'_{AB} \\ \phi'_{TR} & \phi'_{TG} & \phi'_{TB} \\ \phi'_{DR} & \phi'_{DG} & \phi'_{DB} \end{bmatrix} \begin{bmatrix} \Delta Q_R \\ \Delta Q_G \\ \Delta Q_B \end{bmatrix} \tag{11b}$$

where $\phi'_{AR} = \partial \phi_A / \partial Q_R$, etc. Since the receptor primaries are a linear transformation of the color mixture primaries, Q_R, Q_G, and Q_B are a linear transformation of the color mixture tri-stimulus values. Thus, recognizing that A, T, and D must be perturbations in the sensory components and treating X, Y, and Z as perturbations in the tristimulus values, Eq. (11b) is identical to Eq. (1)*

The analysis leading to Eq. (11) makes explict that the first operation of the Guth-Ingling vector model, the linear transformation of the color mixture primaries, is equivalent to employing the first-order approximation of Eq. (10) for absolute and difference threshold experiments. That is, the first operation of the Guth-Ingling vector model is valid if small stimulus differences effect linearization of visual processing and if the Guth-Ingling model is restricted only to those experiments that depend upon such small stimulus differences. Thus, this analysis would lead to the conclusion that color-naming, brightness matching, magnitude estimation, and other Brindley class B experiments[19] are formally excluded from the domain of the Guth-Ingling vector model. The matrix coefficients of Eq. (1) are seen to be determined by the derivatives of the nonlinear functions, ϕ, at the point on the function determined by the reference stimulus, Q_0. Changes in the intensity or spectral composition of the reference stimulus necessarily will produce changes in Q_0, consequently there will be corresponding changes in the values of the derivatives, i.e. changes in the matrix coefficients.

---

* In the case of absolute thresholds, Q_0 is zero so $\Delta Q = Q_1$ and $(\Delta X, \Delta Y, \Delta Z) = (X_1, Y_1, Z_1)$.
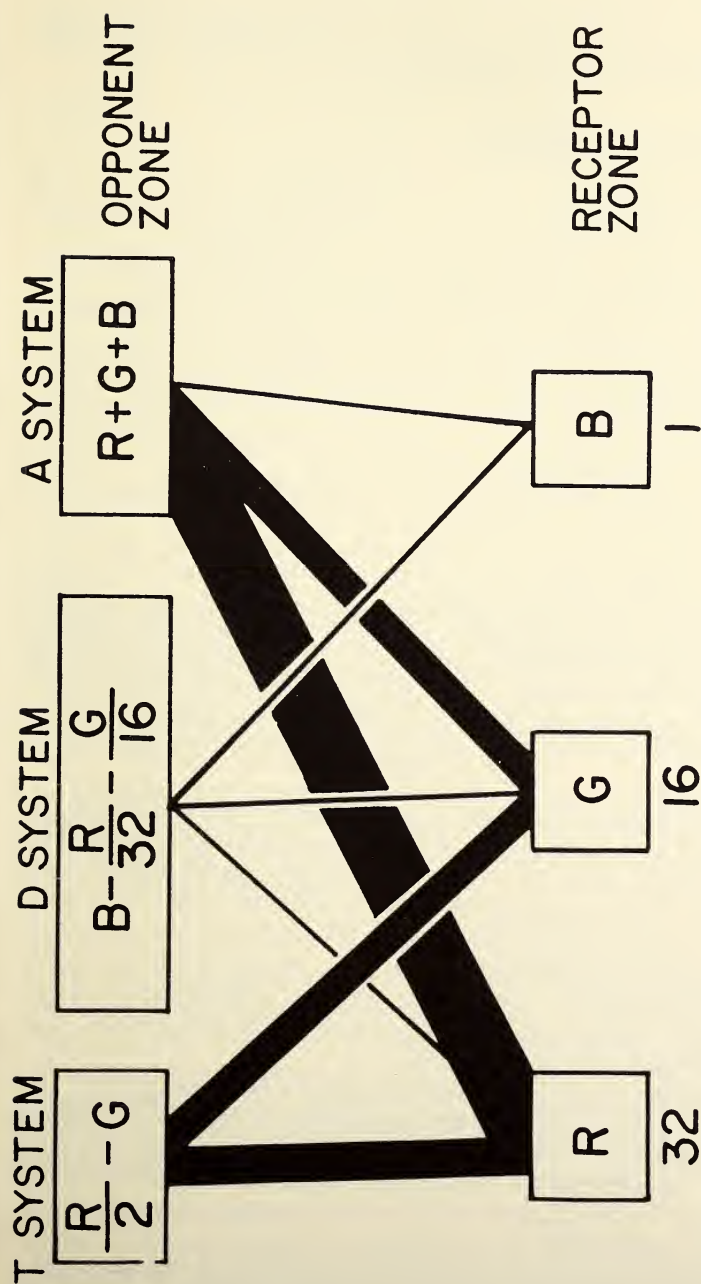
(RELATIVE RECEPTOR WEIGHTS)

Figure 2. Block "wiring diagram" that illustrates physiological interpretation of the *A, T, D* vector model (from Massof, 1977).

Thus, the practice of altering matrix coefficients to account for different t.v.λ. functions, as done by Guth, Ingling and their colleagues,[3-4] is valid in principle (although at present unconstrained).

Vector Magnitude as Detectability

As part of his derivation of the vector space from additivity failure data, Guth defined the second operation of the Guth-Ingling vector model, Eq. (2), to correspond to stimulus detectability. He made explicit that vectors of equal magnitude corresponded to equally detectable stimuli (e.g. $L_{ij} = 1$ in eq. (4a)). Consequently, the definition of vector magnitude as detectability was made an inherent part of the original vector space.

Kranda and King-Smith,[6] in the presentation of their version of a nearly identical vector model of color vision, suggested that the vector magnitude operation of Eq. (2) was a special case of probability summation and could be derived from an empirical approximation presented by Quick.[20] Quick's approximation to probability summation, as applied to the Guth-Ingling vector model, is

$$d = \left[A^{p} + \mid T \mid^{p} + \mid D \mid^{p} \right]^{\frac{1}{p}} \tag{12}$$

where p is determined from the slope of the frequency-of-seeing function. Kranda and King-Smith employed a value of $p = 3.95$ for their predictions. Although Quick's approximation to probability summation is reasonable when following his arguments, the idea of probability summation as a basis for detection, particularly for tasks other than absolute threshold, is unreasonable.

Using the notation of Eqs. (8)-(11), probability summation entails the assumption that V_{A}, V_{T}, and V_{D} are statistically independent and that all variability in sensation is associated with V_{1}; there is no variability in V_{0}. Thus, probability summation assumes that the observer would never mistakenly report detecting a difference between the test and reference stimuli when there is no physical difference, i.e. $Q_{1} = Q_{0}$. Although it is reasonable to assume statistical independence of the sensory components, it is unreasonable to assume that variability only is associated with the test stimulus and not with the reference stimulus.

With the notable exceptions of Buchsbaum and his colleagues[21-23] and Wandell[24], there has been relatively little application of the theory of signal detection to problems in color vision. Probability summation is a very specific detection model that can be derived from the more general theory of signal detection.[25] Although one could postulate specific alternative models to probability summation, it is preferable to return to first principles and monitor the introduction of assumptions required for the derivation of Eq. (2). In this way we will have full knowledge of what the vector

model entails and we will better understand the sources of its success and the limitations on its performance.

In deriving a general formulation for describing an observer's performance in a visual discrimination or detection task, we begin with the assumption that there are random fluctuations in the visual sensation components (V_A, V_T, V_D). These sensory fluctuations could be dependent upon the visual system state that is determined by Q. Thus, most generally, the fluctuations in the visual sensation vector, V, will be represented by a joint-conditional probability density function, f (V | Q).

The specific form of f (V | Q) is unknown, however, in most instances a Gaussian density function willserve as a good approximation. The Gaussian approximation follows from the central limit theorem, which states that the sums of independent random variables (irrespective of the distributions) will approximate a Gaussian distribution.* Employing the Gaussian approximation then,

$$f(V \mid Q_i) = (2\pi)^{-\frac{n}{2}} \mid C_i \mid^{-\frac{1}{2}} \exp{-\frac{1}{2}} \left[(V' - \overline{V}'_i) C_i^{-1} (V - \overline{V}_i) \right] \quad (13)$$

where n is the number of components in the visual sensation vector ($n = 3$ for the Guth-Ingling vector model), $\mid C_i \mid$ is the determinant of the covariance matrix for stimulus condition i (determined by Q_i), C_i^{-1} is the inverse of the covariance matrix, V' is the transpose of the sensory vector V and $\overline{V}_i$ is the mean value of the sensory vector V, given stimulus condition i. The covariance matrix includes the variances of each sensory vector component (along the diagonal) and the covariances between the sensory vector components (in the off-diagonal cells).

Observers can employ any of a number of rules in deciding what their response will be to a stimulus. Some rules are irrational, some are rational but non-optimal, and some are optimal. All optimal rules are equivalent to the likelihood ratio test,

$$l(V) = \frac{f(V \mid Q_1)}{f(V \mid Q_0)} \geq \beta \quad (14)$$

which says, respond "stimulus 1" (e.g. test) if the likelihood ratio is greater than or equal to some criterion β and respond "stimulus 0" (e.g. reference) if the likelihood ratio is less than β. The likihood ratio test determines the performance boundaries, i.e. the best possible performance, or the worst possible performance if the responses are reversed. Since we do not know the observer's actual decision rule, we will assume that it is optimal, and therefore equivalent to the likelihood ratio test of Eq. (14).

---

* We might consider φ to be decomposable into many suboperations that have associated random fluctuations. If the fluctuations are relatively small then linearization might be effected and superposition could lead to Gaussian-limited behavior.

Robert W. Massof
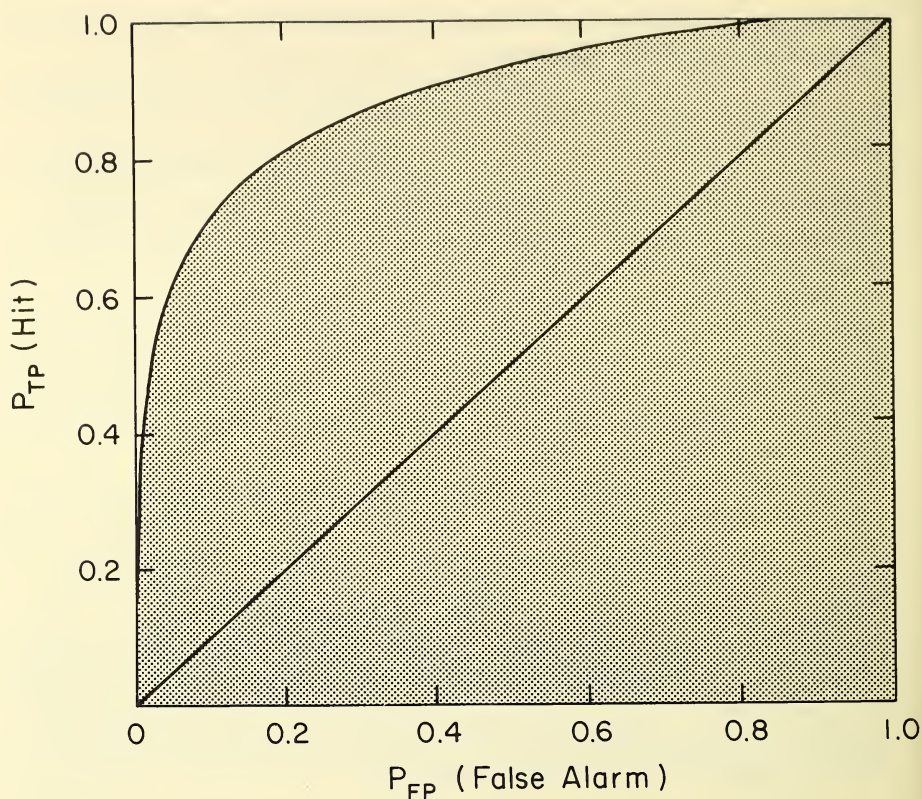
Figure 3. The ROC curve (solid curve) is a plot of the covariance of the probability of making a true-positive response (i.e. responding Q_1 when Q_1 was presented) with the probability of making a false-positive response (i.e. responding Q_1 when Q_0 was presented). The area of the ROC curve (shaded area) is equal to the probability correct in a two-alternative forced-choice detection or discrimiantion task. The area of the ROC curve is a nonparametric index of detectability or discriminability.

Before continuing with the derivation of the detectability definition of Eq. (2), let us reiterate the new assumptions that have been introduced. First, we assumed that given static and uniform stimulation there are random fluctuations in the visual sensation vector, V, which may depend upon the stimulus conditions; this assumption is made explicit in the notion of the joint conditional probability density function, $f(V \mid Q)$. Second, given our lack of specific knowledge of the form of $f(V \mid Q)$, we have accepted the Gaussian approximation, made explicit in Eq. (13); the Gaussian approximation was accepted on the strength of the central limit theorem. Third, we have assumed that the observer's decision rule is optimal and therefore equivalent to the likelihood ratio test of Eq. (14).

The likelihood ratio in eq. (14) will have an associated conditional probability density function, $g[l(V) \mid Q]$. Given some criterion β, the probability of responding Q_1 when Q_1 is presented (true-positive or "hit" rate) is

$$P_{TP} = \int_{\beta}^{+\infty} g[l(V) \mid Q_1] dl(V) \qquad (15)$$

and the probability of responding Q_1 when Q_0 is presented (false-positive or "false-alarm" rate) is

$$P_{FP} = \int_{\beta}^{+\infty} g[l(V) \mid Q_0] dl(V) \qquad (16)$$

By varying the decision criterion β from zero to infinity, a relative operating characteristic (ROC) curve can be constructed. As shown in fig. 3, the ROC curve is a plot of the covariance of P_{TP} and P_{FP} as a function of β. the area of the ROC curve is equal to the probability correct in a two alternative force-choice task and is a nonparametric index of the discriminability of Q_1, from Q_0.[25]

If the density functions that determine the ROC curve are Gaussian (for our example, the likelihood ratio density functions, $g[l(V) \mid Q]$ are not Gaussian), then P_{TP} and P_{FP} can be transformed to z-scores, $Z_{TP} = (x - \mu_1)/\sigma_1$ and $Z_{FP} = (x - \mu_0)/\sigma_0$. Since both z-scores are linear with x, the ROC curve will plot as a straight line

$$z_{TP} = \frac{\sigma_0}{\sigma_1}\left[z_{FP} - \frac{(\mu_1 - \mu_0)}{\sigma_0}\right] \qquad (17a)$$

If $\sigma_1 = \sigma_0$, then

$$z_{TP} = z_{FP} - d' \qquad (17b)$$

where $d' = (\mu_1 - \mu_0)/\sigma_0$. Under these highly constrained conditions, d' completely specifies the ROC curve and therefore is an index of discriminability between the density functions.

The z-score for the area of the binormal ROC curve is

$$z_A = \frac{\mu_1 - \mu_0}{[\sigma_1^2 + \sigma_0^2]^{1/2}} \qquad (18)$$

which is the distance of the normal-deviate ROC from the origin along the perpendicular to the ROC. For the equal variance conditions, $\sigma_1 = \sigma_0$, eq. (18) reduces to

$$z_A = \frac{d'}{\sqrt{2}}. \tag{19}$$

Consequently, z_A is equivalent to d'. However, z_A is more general than d^1 because no assumptions are made about σ_1 and σ_0.

In order to employ eq. (18), it is necessary to transform l (V) to a Gaussian-distributed variable. If the transformation is monotonic, decisions based on the new variable will also be optimal and the ROC curve will be generated. To develop the argument, consider first the Gaussian scalar variable x, conditionally distributed as

$$f(x \mid h_i) = \frac{1}{\sigma_i \sqrt{2\pi}} e^{-1/2\frac{(x-\mu_i)^2}{\sigma_i^2}} \tag{20}$$

The likelihood ratio is

$$l(x) = \frac{f(x \mid h_1)}{f(x \mid h_0)} = \frac{\sigma_0 e^{-1/2\frac{(x-\mu_1)^2}{\sigma_1^2}}}{\sigma_1 e^{-1/2\frac{(x-\mu_0)^2}{\sigma_0^2}}} \tag{21}$$

and the logarithm of the likelihood ratio is

$$L(x) = \ln l(x) = \ln\frac{\sigma_0}{\sigma_1} + \frac{1}{2}\frac{(x-\mu_0)^2}{\sigma_0^2} - \frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}. \tag{22}$$

The log likelihood ratio is a monotonic function of the likelihood ratio, therefore it is an equivalent decision variable. Thus, the ROC curved based on $L(x)$ will be identical to the ROC curve based on $l(x)$.

Equation (22) can be expanded in a Taylor series about μ_1

$$L(x) = L(\mu_1) + \frac{dL(\mu_1)}{dx}(x-\mu_1) + \frac{d^2L(\mu_1)}{dx^2}\frac{(x-\mu_1)^2}{2} \tag{23}$$

or

$$L(x) = \ln\frac{\sigma_0}{\sigma_1} + \frac{1}{2}\frac{(\mu_1-\mu_0)^2}{\sigma_0^2} + \left[\frac{\mu_1-\mu_0}{\sigma_0^2}\right](x-\mu_1) + \left[\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right]\frac{(x-\mu_1)^2}{2} \tag{24}$$

The first-order term in eq. (24) is linear with x. Since linear functions of Gaussian variables are Gaussian, the first-order term in eq. (24) will be Gaussian-distributed. The second-order term in eq. (24) will vanish if $\sigma_1 = \sigma_0$, and will be insignificant unless the difference between σ_1^2 and σ_0^2 is large. Consequently, to the first-order, $L(x)$ will be Gaussian-distributed.

Following exactly the same reasoning from eq. (13) and (14), the log likelihood ratio for the visual sensation vector, $L(V)$, will be approximately Gaussian-distributed since the first-order term is a linear function of V. The magnitude of the second-order term is proportional to the difference between C_0^{-1} and C_1^{-1} (re. eq. 13). Thus, if $C_0 = C_1$ the second-order term will vanish and $L(V)$ will be Gaussian. Under these first-order assumptions, eq. (18) becomes

$$z_A = \frac{E[L(V) \mid Q_1] - E[L(V) \mid Q_0]}{\sqrt{VAR[L(V) \mid Q_1] + VAR[L(V) \mid Q_0]}} \tag{25}$$

where $E[L(V) \mid Q_i]$ is the expected value of the log likelihood ratio given condition i and $VAR[L(V) \mid Q_1]$ is the corresponding variance.

The detectability expression for the Guth-Ingling vector model, Eq. (2), follows from Eq. (25) if we assume that the covariance matrices, C_i in Eq. (13), are equal and diagonal.[26] That is, $C_1 = C_0$ and the sensory vector components are statistically independent, i.e. all covariances are zero. To see this, from Eqs. (13) and (14), the log likelihood ratio is

$$L(V) = \frac{1}{2}(V' - \overline{V}'_0)C_0^{-1}(V - \overline{V}_0) - \frac{1}{2}(V' - \overline{V}'_1)C_1^{-1}(V - \overline{V}_1) + \frac{1}{2}\ln\frac{|C_0|}{|C_1|} \tag{26}$$

From the statistical independence assumption,

$$C_0^{-1} = \begin{bmatrix} \sigma_{A_0}^{-2} & 0 & 0 \\ 0 & \sigma_{T_0}^{-2} & 0 \\ 0 & 0 & \sigma_{D_0}^{-2} \end{bmatrix} \tag{27}$$

and from the assumption that $C_1 = C_0$, Eq. (26) becomes

$$L(V) = \frac{1}{2}\left[\frac{\left(\overline{V}_{A_0}^2 - 2V_A \overline{V}_{A_0} + 2V_A \overline{V}_{A_1} - \overline{V}_{A_1}^2 \right)}{\sigma_{A_0}^2} \right. \tag{28}$$
$$+ \frac{\left(\overline{V}_{T_0}^2 - 2V_T \overline{V}_{T_0} + 2V_T \overline{V}_{T_1} - \overline{V}_{T_1}^2 \right)}{\sigma_{T_0}^2}$$
$$\left. + \frac{\left(\overline{V}_{D_0}^2 - 2V_D \overline{V}_{D_0} + 2V_D \overline{V}_{D_1} - \overline{V}_{D_1}^2 \right)}{\sigma_{D_0}^2} \right]$$

Taking the expected values required in Eq. (25), we obtain

$$E\left[L(V)\mid Q_0\right] = \frac{1}{2}\left[\frac{\left(-\bar{V}_{A_0}{}^2 + 2\bar{V}_{A_0}\bar{V}_{A_1} - \bar{V}_{A_1}{}^2\right)}{\sigma_{A_0}{}^2}\right. \tag{29a}$$
$$+ \frac{\left(-\bar{V}_{T_0}{}^2 + 2\bar{V}_{T_0}\bar{V}_{T_1} - \bar{V}_{T_1}{}^2\right)}{\sigma_{T_0}{}^2}$$
$$\left.+ \frac{\left(-\bar{V}_{D_0}{}^2 + 2\bar{V}_{D_0}\bar{V}_{D_1} - \bar{V}_{D_1}{}^2\right)}{\sigma_{D_0}{}^2}\right]$$

and

$$E\left[L(V)\mid Q_1\right] = \frac{1}{2}\left[\frac{\left(\bar{V}_{A_1}{}^2 - 2\bar{V}_{A_1}\bar{V}_{A_0} + \bar{V}_{A_0}{}^2\right)}{\sigma_{A_0}{}^2}\right. \tag{29b}$$
$$+ \frac{\left(\bar{V}_{T_1}{}^2 - 2\bar{V}_{T_1}\bar{V}_{T_0} + \bar{V}_{T_0}{}^2\right)}{\sigma_{T_0}{}^2}$$
$$\left.+ \frac{\left(\bar{V}_{D_1}{}^2 - 2\bar{V}_{D_1}\bar{V}_{D_0} + \bar{V}_{D_0}{}^2\right)}{\sigma_{D_0}{}^2}\right]$$

In terms of expected values, the variance of the log likelihood ratio distribution, given Q_i, is

$$VAR\left[L(V)\mid Q_i\right] = E\left[L(V)^2\mid Q_i\right] - E\left[L(V)\mid Q_i\right]^2 \tag{30}$$

which under assumption of equal and diagonal covariance matrices (i.e. Eq. (27) and $C_1 = C_0$) leads to

$$VAR\left[L(V)\mid Q_1\right] = VAR\left[L(V)\mid Q_0\right] = \frac{(\bar{V}_{A_1}-\bar{V}_{A_0})^2}{\sigma_{A_0}{}^2} + \frac{(\bar{V}_{T_1}-\bar{V}_{T_0})^2}{\sigma_{T_0}{}^2} + \frac{(\bar{V}_{D_1}-\bar{V}_{D_0})^2}{\sigma_{D_0}{}^2} \tag{31}$$

Substituting Eqs. (29a), (29b), and (31) into Eq. (25) we obtain

$$z_A = \left[\frac{(\bar{V}_{A_1}-\bar{V}_{A_0})^2}{2\sigma_{A_0}{}^2} + \frac{(\bar{V}_{T_1}-\bar{V}_{T_0})^2}{2\sigma_{T_0}{}^2} + \frac{(\bar{V}_{D_1}-\bar{V}_{D_0})^2}{2\sigma_{D_0}{}^2}\right]^{1/2} \tag{32a}$$

which from Eqs. (9a) and (19) reduces to

$$d' = \left[\frac{\Delta V_A^2}{\sigma_{A_0}{}^2} + \frac{\Delta V_T^2}{\sigma_{T_0}{}^2} + \frac{\Delta V_D^2}{\sigma_{D_0}{}^2}\right]^{1/2} \tag{32b}$$

Equation (32b) is identical to Eq. (2), the second operation of the Guth-Ingling vector model, if the standard deviations $(\sigma_{A_0}, \sigma_{T_0}, \sigma_{D_0})$ are regarded to be implicit in the transformation coefficients of Eq. (1).*

It is important to remember that two specific assumptions were required to achieve Eq. (32): 1) the covariance matrices are equal and 2) the visual sensation vector components are statistically independent. The assumption of equal covariance matrices is equivalent to assuming that variability in sensation arises from an added noise source. The assumption that the sensory vector components are independent is implicit in Guth's use of the factor analysis in his derivation of the vector model from heterochromatic luminance additivity failure data.[4]

Limitation on Applications of the Guth-Ingling Vector Model

The assumptions required for the derivation of the Guth-Ingling vector model must be regarded as first-order approximations. For example, the assumption of equal convariance matrices is in conflict with the conclusions drawn from past studies of vision ROC curves.[28-30] However, we may regard the equal covariance matrices assumption to be an approximation which would be equivalent to replacing both C_0 and C_1 in Eq. (26) with the average matrix. Also, to the extent that the R, G, B cones contribute to all three sensory components, and considering global physiological variations that may affect the entire visual system, it is unlikely that the three sensory components are statistically independent. Yet, we can accept the assumption of statistical independence among sensory components if we regard that assumption to be equivalent to a transformation of the components that would diagonalize the covariance matrices (that is, the covariance terms would be implicit in the derived matrix coefficients, ϕ^1_{ij} in Eq. (11b)). The other assumptions, viz. linearization, Gaussian approximation, and likelihood ratio observer also are likely to fail under close scrutiny; nevertheless, they are acceptable as first-order descriptions.

The predictions of the Guth-Ingling vector model typically are compared to data that are averaged across several observers. Such data usually are means of average individual observations. It probably is safe to assume that for such data between-observer variability will outweigh within-observer variability. If so, individual departures from the first-order assumptions required for the derivation of the Guth-Ingling vector model would be averaged out. This "averaging-out" of higher-order terms would lead to an improved performance of the first-order assumptions. Thus, the

---

* Eq. (32) also follows from Prucnal's general performance parameter, given the same assumptions.[27]
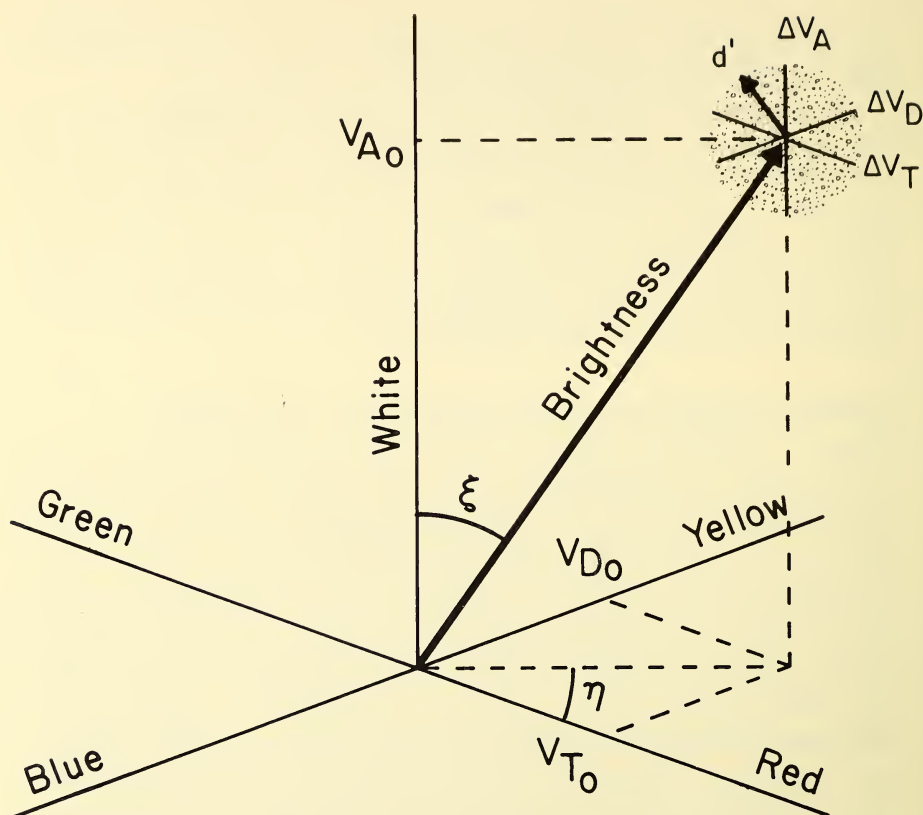
Figure 4. Visual sensation geometry with axes representing Hering elemental color sensations: white (V_A), red or green (V_T), and blue or yellow (V_D). The reference stimulus produces the visual sensation represented by the origin-bound vector. The saturation of the reference stimulus is represented by the angle ξ, hue is represented by the angle η, and brightness is represented by the magnitude of the vector. The reference stimulus (head of the vector) defines the origin of a subspace (set of coordinates at the head of the vector) representing perturbations in visual sensation. Within limits (represented by the sphere), distances in the subspace can be computed from basis vectors that are approximated by linear transformations of stimulus perturbations. Distance from the subspace origin corresponds to d'. Increment thresholds are represented by d' vectors above the ΔV_T, ΔV_D plan and decrement thresholds are represented by d' vectors below the ΔV_T, ΔV_T plane. Color discriminations are represented by d' vectors in the ΔV_T, ΔV_T plane.

Guth-Ingling vector model is likely to rest on a theoretically sound foundation when applied to the threshold and discrimination data of the "average observer."

To understand the limitations on the Guth-Ingling vector model it is useful to employ graphical representations of the above analysis. Granted the assumptions leading to Eqs. (11b) and (32b), the Guth-Ingling vector model can be thought of as a subspace of a geometry for visual sensation, a geometry that is built on the visual sensory components, V_A, V_T, V_D. Figure 4 graphically illustrates this concept.

The "red-green" opponent colors are represented by opposite poles of one bipolar dimension (V_T), the "blue-yellow" opponent colors are represented by opposite poles of a second bipolar dimension (V_D), and "white" is represented by a third monopolar dimension (V_A) (we will ignore "black" in this presentation). Within this scheme, the visual sensation produced by the reference stimulus is represented by a visual sensation vector, V_0. Vector magnitude is a monotonic function of brightness, and the two angles that specify vector direction correspond to the components of color; the polar angle ξ is a monotonic function of saturation and the azimuthal angle η is a monotonic function of hue.

The visual sensation produced by the reference stimulus corresponds to a point in this 3-space that is determined by V_{A0}, V_{T0}, and V_{D0} (the head of the vector). This point in turn, would define the origin of a subspace. The axes of the subspace would be parallel to the sensory axes and would represent perturbations of the sensory components, ΔV_A, ΔV_T, ΔV_D. The detectability of the sensor perturbation, d', would correspond to the magnitude of an origin-bound vector in this subspace. The boundaries of the subspace (represented by the sphere in fig. 4) would represent the limits of linearization. That is, all points within the sphere could be determined, within acceptable limits, by the linearization. That is, all points within the sphere could be determined, within acceptable limits, by the linear equations of Eq. (11b).

The complete Guth-Ingling subspace would provide for both negative and positive values on ΔV_A, although the white sensory component, ΔV_A, may be always positive. Points within the sphere below the ΔV_T, ΔV_D plane would correspond to decrement thresholds, points above the ΔV_T, ΔV_D plane would correspond to increment thresholds. Points on the ΔV_T, ΔV_D plan would correspond to color discrimination thresholds about the reference stimulus and points along the principal ray through the subspace origin, which may be considered an extension of the vector representing the reference stimulus sensation, would correspond to homochromatic brightness discriminations. Because of the limits on linearization, the explanatory domain of the Guth-Ingling vector model is restricted to the volume of the Guth-Ingling subspace. That is, the Guth-Ingling vector

model cannot legitimately be applied to experiments that entail points outside of the bounded subspace. This limitation is not a property of the geometry, it is a property of the nonlinear functions, ϕ in Eq. (7) and (8), that characterize the visual system's transformation of light absorption by the visual pigments to visual sensation.

Brightness Matching and Luminance

Guth, Ingling and their colleagues defined brightness as vector magnitude in their models for the purpose of generating predictions of heterochromantic brightness matching spectral sensitivity and generating predictions of brightness-(and lightness)-to-luminance ratios.[4] The definition of brightness necessarily entails the strong, and erroneous assumption that brightness is linear with stimulus intensity. The success of the Guth-Ingling vector model in accounting for heterochromatic brightness-matching spectral sensitivity is a consequence of the coincidental similarity between brightness-matching spectral sensitivity data and absolute threshold spectral sensitivity data.[2] The Guth-Ingling vector model legitimately accounts for only the absolute threshold spectral sensitivity data. Because heterochromatic brightness matching involves large sensory differences between the reference and comparison stimuli (they are different colors), linearization fails; consequently, heterochromatic brightness matching measures fall outside the domain of the Guth-Ingling vector model.

Guth and Ingling both defined the A component of their models to be equal to luminance. In terms of Eq. (11b), this means that $\Delta V_A = a_{12}\Delta Y$. This is, irrespective of the reference stimulus, ΔX and ΔZ do not contribute to ΔV_A (however, a_{12} will depend on Q_0). Since this definition is in keeping with the structure of the vector model, it is theoretically legitimate, although arbitrary. As for spectral sensitivity, predictions of spectral brightness-to-luminance ratios are a property of the similarity between absolute threshold spectral sensitivity data and brightness matching spectral sensitivity data, not a property of the model. Failure of linearization makes such prediction illegitimate. The vector model predictions of non-spectral lightness-to-luminance ratios (using brightness-to-luminance equations), which are only qualitatively successful at best, also are illegitimate since linearization fails under these experimental conditions as well.

Spectral Hue-Names and Saturation

Guth, Massof and Benzschawel[4] generated predictions of apparent hue and apparent saturation as a function of wavelength. These predictions were qualitatively similar to the spectral color-naming data of Boynton and Gordon[31] and the saturation scaling data of Indow and Stevens[32]. The color-

naming predictions from the Guth-Ingling vector model were generated using the equations.

$$\% \ Red \ or \ Green = \frac{|T_\lambda|}{|T_\lambda| + w \ |D_\lambda|} \tag{33}$$

and

$$\% \ Blue \ or \ Yellow = \frac{w \ |D_\lambda|}{|T_\lambda| + w \ |D_\lambda|} \tag{34}$$

where the weight, w, on $D\lambda$ was set to 4 or 12, depending on the data set used for comparison. The saturation scaling predictions were generated from the equation

$$Saturation = \frac{Chromatic}{Chromatic + Achromatic} = \frac{|T_\lambda| + |D_\lambda|}{A_\lambda + |T_\lambda| + |D_\lambda|}. \tag{35}$$

Different saturation curves were generated by adjusting relative weights on A, T, and D.

For both color-naming and saturation scaling, the observer is asked to assign a value to the visual sensation produced by the stimulus. For this experimental paradigm, Eq. (7) is the appropriate description of the transformation of the static and uniform stimulus into the static and uniform components of visual sensation. Extracting color names or saturation scales involves further observer operations on the visual sensation components. However, since the initial transformation entails nonlinear operations, color-naming and saturation scaling data do not meet the linearization assumption required by the Guth-Ingling vector model.

In graphical terms, each stimulus produces a different vector in the V_A, V_T, V_D space of fig. 4 The observer is instructed to assign a value to the visual sensation that corresponds to the angle η in the case of color-naming, or to the angle ξ in the case of saturation scaling. Since no linearized subspaces are involved in the representation of these scaling tasks, such experiments fall outside the domain of the Guth-Ingling vector model.

Wavelength Discrimination, Saturation Discrimination, and Chromaticity Discrimination

Using the Guth-Ingling vector model, Ingling and Tsou[4] and Guth, Massof, and Benzschawel[3] published predictions of wavelength discrimination and saturation discrimination; Guth *et al* also published predictions of chromaticity discrimination. Such discriminations entail just-noticeable-differences between the test and reference stimuli, linearization ought to be effected and therefore wavelength discrimination, saturation discrimination, and chromaticity discrimination ought to fall within the domain of

Robert W. Massof
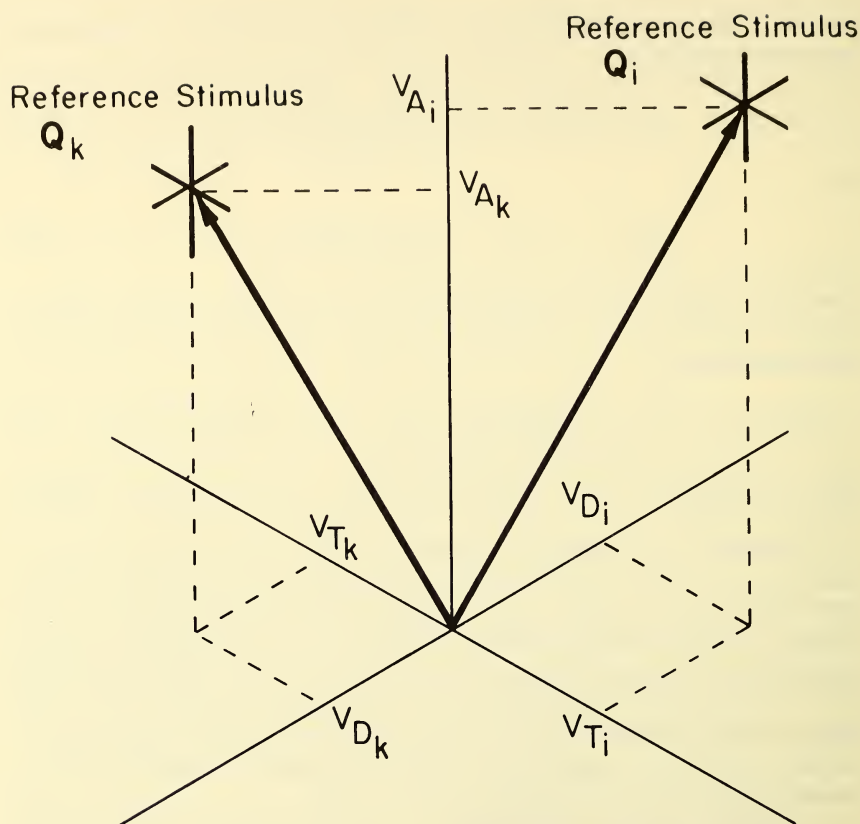
Figure 5. Same as figure 4 but illustrating two different Guth-Ingling subspaces corresponding to two different reference stimuli (Q_i and Q_k). In the case of chromaticity-discrimination ellipses, each ellipse is centered on a different reference stimulus, therefore, each ellipse would be computed from a different subspace. Owing to visual system non-linearities, each subspace would entail a different system of linear equations to transform stimulus perturbations into sensory perturbations.

the Guth-Ingling vector model. Although such predictions are legitimate within the framework of the theory, Ingling, Guth and their colleagues employed improper operations on the vector model in their attempts to account for color discrimination data. At issue is the way the reference stimulus is represented within the vector model.

Although different specific equations were used to account for different data sets, the same general rule was used by Ingling and Guth for all color discrimination predictions. Two vectors were generated in the A, T, D space, one for the reference stimulus and one for the comparison, or test stimulus. If the heads of the vectors were separated by a criterion distance, they were considered to represent discriminably different stimuli. Changes in the reference stimulus, in the case of chromaticity discrimination and wavelength discrimination, was handled by generating a new reference stimulus vector within A, T, D space and determining the value of the test stimulus that corresponded to the vector that was a criterion distance away.

From the theoretical derivation of the Guth-Ingling vector model, the reference stimulus would always correspond to the origin of the linear subspace (i.e. V_{A0}, V_{T0}, V_{D0}). So, as illustrated in figure 5, for chromaticity discrimination there would be a different vector in V_A, V_T, V_D space representing each reference stimulus of different chromaticity. A Guth-Ingling subspace would be defined for each reference stimulus, with its origin at the head of the respective reference sensation vector. Color discrimination would then be computed as distances from the origin of the subspace. That is, detectability of a just-noticeable difference is always represented by the magnitude of an origin-bound vector. Analytically, this explanation means that new matrix coefficients, ϕ'_{ij} in Eq. (11b), would have to be computed for each reference stimulus wavelength in wavelength discrimination and for each reference stimulus chromaticity in chromaticity discrimination. This conclusion follows from linearization of Eq. (9b) for different values of Q_0 corresponding to different reference stimuli.

In summary, the Guth-Ingling vector model is formally limited to generating predictions of data from threshold and discrimination tasks that require just-noticeable-differences in sensation. Such tasks entail small sensory perturbations that are likely to ensure linearization. In contrast, magnitude estimation, color scaling, and experiments that employ large differences between comparison and test stimuli (e.g. heterochromatic brightness matching) are formally excluded from the explanatory domain of the Guth-Ingling vector model. In the case of threshold and discrimination measures, linearization occurs around the state of the visual system determined by the reference stimulus. Therefore, a new set of linear coefficients in the vector model will have to be computed for each reference stimulus. The assumptions required for the derivation of vector magnitude as representing detectability are first-order approximations and most likely

are satisfied when considering the pooled data of several observers. Thus, application of the Guth-Ingling vector model would be most effective for the "average observer."

Discussion

The analysis presented here shows that the Guth-Ingling vector model implicitly rests on a firm theoretical foundation. The specific assumptions required for the derivation of the Guth-Ingling vector model are first-order approximations that most likely are satisfied by the "average observer." To the extent that experimental conditions satisfy the static and uniform constraint, and the detection or discrimination task entails small sensory differences that would effect linearization, application of the Guth-Ingling vector model is appropriate. Indeed, the Guth-Ingling vector model enjoys strong theoretical leverage because the assumptions are explicit and reasonable. Superior models to the Guth-Ingling vector model would have to entail high-order terms in the theory, such as more specific and detailed covariance matrices, or precise definitions of non-Gaussian sensory density functions, or higher order terms in the Taylor series expansions of the nonlinear functions, etc.

This paper offers only the structure and interpretation of a color detection and discrimination model, it does not offer the details. For application to psychophysical data, it will be necessary to invent specific nonlinear functions for ϕ_A, ϕ_T, and ϕ_D in Eqs. (8a)-(8c). Once these functions are defined, the model then will be constrained to generate predictions of chromaticity discrimination, wavelength discrimination, saturation discrimination, absolute and increment threshold spectral sensitivities, heterochromatic, Stiles-type, threshold-versus-radiance functions, heterochromatic absolute and increment threshold additivity failures, etc., all as a function of reference stimulus luminance and chromaticity. Once the nonlinear functions have been defined, there are no additional degrees of freedom or *ad hoc* operations that can rescue the model when it fails. Only the form of the nonlinear functions or the parameters in the nonlinear functions can be adjusted to improve the performance of the model.

Other operations on the sensory components are required to account for color appearance and brightness, such undefined operations and the first-order assumptions entailed by the Guth-Ingling vector model, make it advisable to extend predictions of the nonlinear functions to data sets outside of the domain of the Guth-Ingling vector model. Such predictions would not be supported by theory.

The next task in developing applications of the general theory is to explore interrelationships among various color detection and discrimination data sets. Spatial and temporal boundaries must be set on the static

and uniform constraint, since different data sets obtained with different spatial and temporal parameters theoretically may not be comparable. Also, color discriminations have long been modeled geometrically through the use of line elements. Through derivations similar to those described in the present paper, line elements can be derived from the general theory.[33] Such derivations provide insight into the nature of metric coefficients of various line elements, the assumptions implied by specific line element models, and the formal connections of line element models to the Guth-Ingling vector models.

References

1. Wyszecki, Gunter and W. S. Stiles, *Color Science*, 2nd ed., Wiley, New York, 1982.
2. Guth, Sherman L. and Howard L., Lodge Heterochromatic additivity, foveal spectral sensitivity, and a new color model, *J. Opt. Soc. Am. 63*, 450-462 (1973).
3. Guth, S. Lee, Massof, Robert W. and Benzschawel, Terry, Vector model for normal and dichromatic color vision, *J. Opt. Soc. Am. 70*, 197-212 (1980).
4. Ingling, Jr., Carl R. and Tsou, Brian H., Orthogonal combinations of three visual channels, *vision Res. 17*, 1075-1082 (1977).
5. Ingling, Jr., Carl R., Luminance and opponent color contributions to visual detection and to temporal and spatial integration, *J. Opt. Soc. Am. 68*, 1143-1146 (1978).
6. Kranda, K. and King-Smith, P. E., Detection of coloured stimuli by independent linear systems, *Vision Res. 19*, 733-746 j(1979).
7. Benzschawel, Terry and Guth, S. Lee ATDN: Toward a uniform color space, *Color Res Appl. 8*, 133-141 (1984).
8. Yaguchi, Hirohisa and Ikeda, Mitsuo, Nonlinear nature of the opponent-color channels, *Color Res. Appl. 7*, 187-190 (1982).
9. Massof, Robert W., Color vision theory and linear models of color vision, *Color Res. Appl. 10*, 133-146 (1985).
10. Guth, Sherman L., Nonadditivity and inhibition among chromatic luminaces at threshold, *Vision Res. 7*, 319-328 (1967).
11. Guth, Sherman L., Donley, Nola J., and Marrocco, R. T., On luminance additivity and related topics, *Vision Res 9*, 537-575 (1969).
12. Swets, J. A., Color vision, *Quarterly Progress Report, Research Laboratory of Electronics*, Mass. Inst. of Tech., Cambridge, MA (1960).
13. Cohen Jozef, and Gibson, W. A., Vector model for color sensations, *J. Opt. Soc. Am. 52*, 692-697 (1962).
14. Guth, Sherman Leon (Lee), A new vector model, in: *Color Metrics* (J.J. Vos, L.F.C. Friele, and P.L. Walraven, Eds.) pp. 82-98, AIC/Holland, Institute for Perception TNO, Soesterberg, 1972.
15. Massof, Robert W., A quantum fluctuation model for foveal color thresholds, *Vision Res. 17*, 565-570 (1977).
16. Vos, J. J. and Walraven, P. L., On the derivation of foveal receptor primaries, *Vision Res. 11*, 799-818 (1971).

17.   Massof, Robert W. and Bird, Joseph F., A general zone theory of color and brightness vision. I. Basic formulation, *J. Opt. Soc. Am. 68*, 1465-1471 (1978).

18.   Bird Joseph F., and Massof, Robert W., A general zone theory of color and brightness vision. II. The space-time field, *J. Opt. Soc. Am. 68*, 1471-1478 (1978).

19.   Brindley, G. S., *Physiology of the Retina and Visual Pathway*, 2nd ed., Williams and Wilkins, Baltimore, 1970.

20.   Quick, Jr., R. F., A vector-magnitude model of contrast detection, *Kybernetic 16*, 65-67 (1974).

21.   Buchsbaum, G. and Goldstein, J. L., Optimum probabilistic processing in colour perception. I. Colour discrimination, *Proc. R. Soc. Lond. 205*, 229-247 (1979).

22.   Buchsbaum, G. and Goldstein, J. L., Optimum probabilistic processing in colour perception. II. Colour vision as template matching, *Proc. R. Soc. Lond. 205*, 249-266 (1979).

23.   Buchsbaum, G. and Gottschalk, A., Trichromacy, opponent colours coding and optimum colour information transmission in the retina, *Proc. R. Soc. Lond. 220*, 89-113 (1983).

24.   Wandell, Brian A., Measurement of small color differences, *Psychol. Rev. 89*, 281-302 (1982).

25.   Green, David M. and Swets, John A., *Signal Detection Theory and Psychophysics*, 2nd ed., Robert E. Krieger, New York, 1974.

26.   Massof, Robert W. and Starr, Stuart J., Vector magnitude operation in color vision models: Derivation from signal detection theory, *J. Opt. Soc. Am. 70*, 870-872 (1980).

27.   Prucnal, Paul R., Generalized performance parameter for single-threshold detection systems, *Appl. Opt. 19*, 3606-3610 (1980).

28.   Nachmias, Jacob and Kocher, Elizabeth C., Visual detection and discrimination of luminance increments, *J. Opt. Soc. Am. 60*, 382-389 (1970).

29.   Cohn, Theodore E., Quantum fluctuation limit in foveal vision, *Vision Res. 16*, 573-579 (1976).

30.   Massof, Robert W., Relation of the normal-deviate vision receiver operating characteristic curve slope to d'. *J. Opt. Soc. Am., A., 4*, 548-550. (1987).

31.   Boynton, Robert M. and Gordon, James, Bezold-Brucke hue shift measured by color-naming technique, *J. Opt. Soc. Am. 55*, 78-86 (1965).

32.   Indow, Tarow and Stevens, S. S., Scaling of saturation and hue, *Percept. Psychophys. 1*, 253-271 (1966).

33.   Massof, Robert W., Nonlinear color vision models applied to color and luminance difference thresholds, *J. Opt. Soc. Am. A 2* (13), p 49, abstract WU4 (1985).

VEILING REFLECTIONS
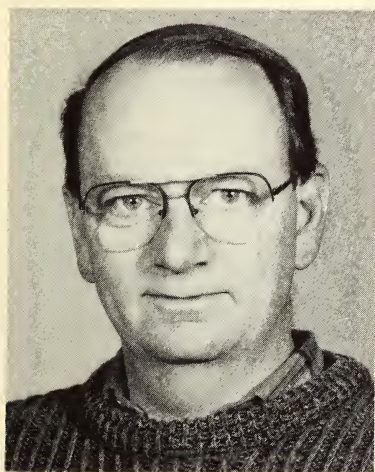
James A. Worthey
Research Scientist at the National Institute of Standards and Technolgy in Gaithersburg, Maryland. He is a member of the Illuminating Engineering Society of North America and the Optical Society of America.

ABSTRACT. Familiar light sources vary in the solid angle they subtend by a factor of 100,000 or more. Large lights hamper a person reading in the attempt to avoid veiling reflections on her book, as the chapter shows in detail. The answer to veiling reflections is to have a dark area at the mirror angle with respect to the eye. A method is developed to quantify veiling reflection amplitude, based on thought experiments in which a black glass sits next to diffusely reflecting white surface. Two conclusions are that veiling reflections are never negligible and that the "dark area" should be dark compared to the illuminated white surface. When a light source is small, veiling reflections are traded for highlights, which can greatly increase the dynamic range of a scene. This leads to the conclusion that a scene diffusely lighted with light of mediocre color rendering looks washed out because it is washed out.

VEILING REFLECTIONS

Introduction

The size of the light source illuminating a scene affects the strength of shading on object surfaces, the sharpness of cast shadows, and the nature of reflections from shiny or glossy surfaces. These fairly obvious effects are not discussed systematically in traditional lighting textbooks. This chapter will attempt an orderly discussion of one segment of these effects: source size and its effects on veiling reflections and highlights.

Familiar light sources vary by a factor of about 10^6 in the bright solid angle they present. Table 1 ranks a number of sources according to the solid angle subtended by their bright areas, based on simple measurements and calculations. For instance, the filament area for the unfrosted 60 W bulb is calculated from the observation that the coiled coil appears to be 1 mm diameter by 20 mm long. Obviously, the choice of two meters distance for the individual electric lights is arbitrary and affects their comparison with the sun and the luminous ceiling. Nevertheless, the range of variation is great.

Table 1, Bright Areas of Light Sources.

Light Source	Bright Area, m²	Solid Angle at 2 m distance, microsteradians
Unfrosted 60 W incand. bulb	2.0×10^{-5}	5
The Sun (distance = 93,000,000 miles)	1.5×10^{18}	67
Ordinary frosted incand. 60 W bulb	3.1×10^{-4}	79
Soft White 60 W incand. bulb	2.4×10^{-3}	590
F40T12 fluorescent tube	$4.6 \times 10_{-2}$	12000
Luminous ceiling, extending to ∞ (2π steradians)	many	6,300,000 (2π million)

From one type of incandescent bulb to the next, bright area increases by a factor of about ten, and the effects of these differences are sufficient that all three types are sold in supermarkets. The transitions to one fluorescent tube and then to the luminous ceiling increase area by further factors of 20 and 500. Practical lighting designers know that small sources give a scene "sparkle" while large ones give veiling reflections and other problems.

How Big is Too Big?

To model the task of avoiding veiling reflections, let us assume that an observer is reading a book on very shiny paper. In fact, let us assume that the book is a flat mirror. To simplify calculations, let us assume a circular light source and a circular book that can be tipped through any angle, about any axis of tipping (Figure 1). I have run a computer simulation of this entire three-dimensional problem, finding just how far the book must be tipped, as a function of the direction in which it is tipped. Running this complicated simulation shows that the book should always be tipped away from the light, with the normal to the book remaining in the plane established by the eyes, the book, and the light. This means that a simpler sketch and calculation are adequate for finding how far to tip the book (Figure 2).

This analysis leads to Equation (1):

$$cd \sin(\psi + 2\beta) - bd \cos(\beta) - cb \cos(\psi + \beta) = 0 \qquad (1)$$

The symbols are defined by Figure 2. While it is not readily possible to solve this equation for β, a numerical solution of it is much simpler than the previous computer simulation.

Figure 3 shows angle of tip versus luminaire size for two sizes of book. The curves are not quite straight. There are no stunning surprises here, just a quantitative statement of the fact that the bigger the light is, the farther you have to tip the book. How big is too big? The answer depends on such factors as the reader's freedom to re-position himself, but Equation (1) is a step toward understanding. More needs to be done in studying veiling reflections by simple examples and the methods of geometrical optics.

How Dark Is Dark Enough?

Although a person will attempt to move the light-source image off of his book, there is always something at the mirror angle with respect to the eyes. That something should be a dark surface. To say how dark this surface must be, we should first understand how veiling reflections arise.

Figure 4 presents a classic textbook picture of pigmented dielectric surfaces. Each colored material is considered to consist of a colorless dielectric medium with colored pigment granules in it. Although this model
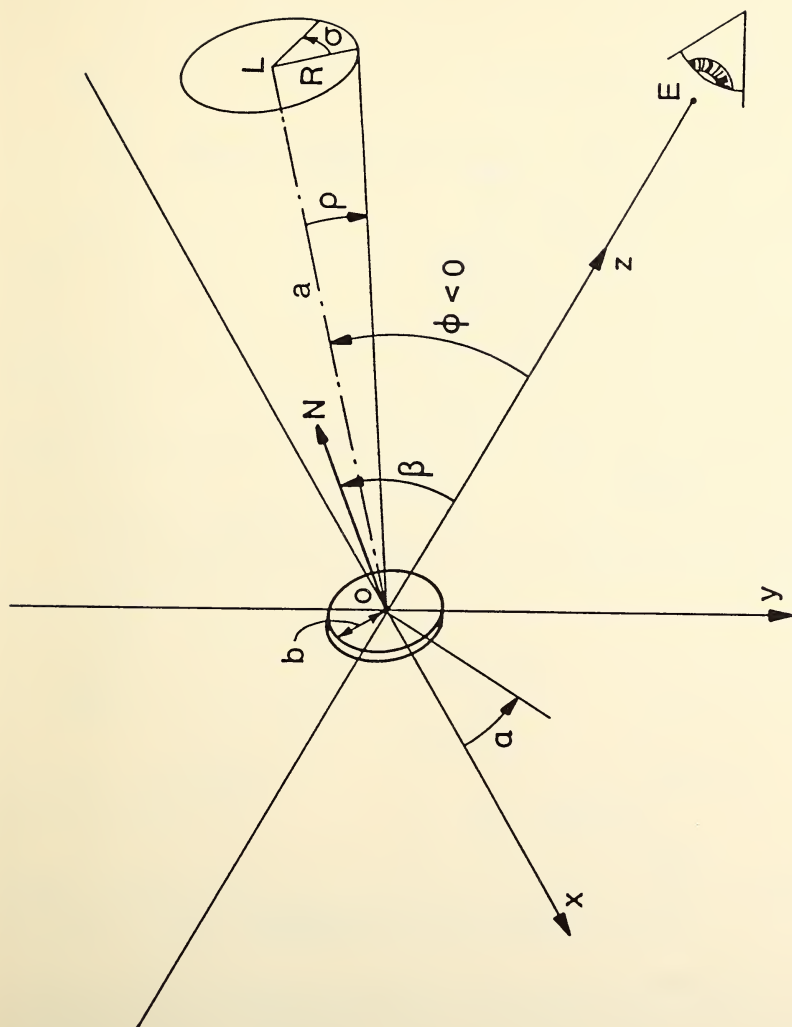
Figure 1. Diagram for three-dimensional analysis of how to tip a book in order to avoid veiling reflections. The book is at the origin and the other circular object is the light source. Detailed solution of this problem showed that a simplified problem statement would suffice.
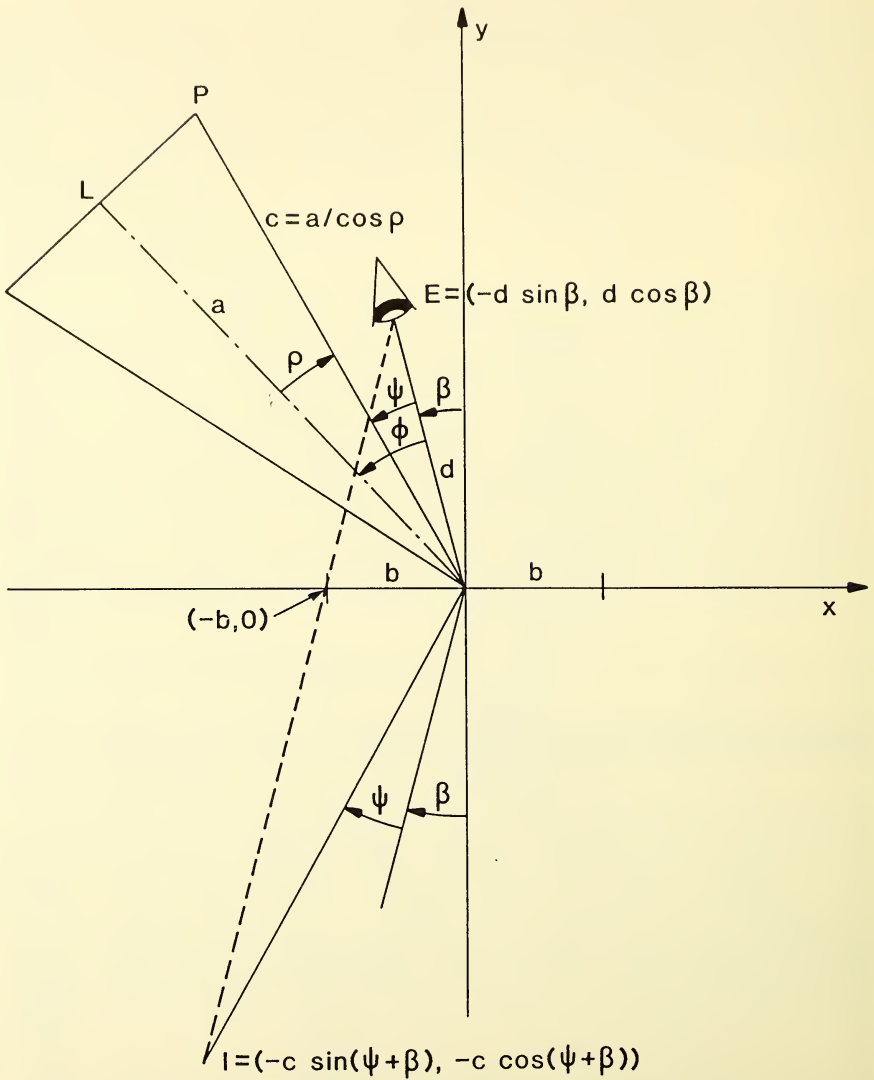
Figure 2. Two-dimensional version of Figure 1. The eye and luminaire are in the *x-y* plane. The book has width 2*b* and the normal to the book coincides with the *y*-axis. Angle ρ is the semi-subtense of the luminaire, while the eye is an angle ϕ from the luminaire center, and an angle β from the normal to the book. The luminaire is a distance *a* from the origin, while the eye is a distance *d*. Two convenient auxiliary variables are ψ and *c*.

Figure 3. Angle of tip β versus the luminaire size parameter ρ, for two values of the book radius *b*. These curves, computed by numerical solution of Equation (1), are not quite straight lines. Here, the luminaire is over the shoulder (φ = 30°).

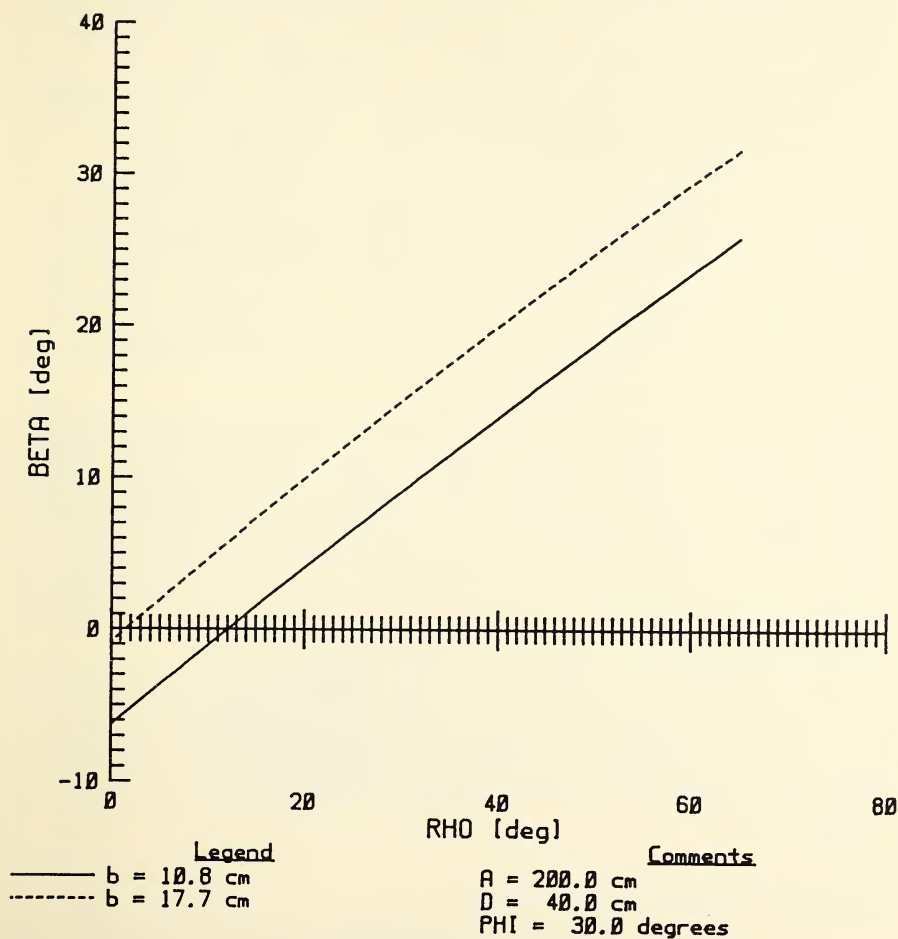is known to be oversimplified, it gives a basic understanding of how light is reflected from most non-metallic objects. A fraction of this incident light is reflected at the surface because of the change in index of refraction there. The exact amount reflected depends on the surface's refractive index, the angle of incidence, and the polarization of the light[1,2,3]. However, if we assume a "textbook" refractive index value of n = 1.50, unpolarized light, and incidence not too far from perpendicular, we can put aside the details and say that the surface reflects four percent of the incident light.

In Figure 4, the more interesting case is the one on the left, the shiny plastic. Because the surface is shiny, the incoming ray gives rise to one outgoing ray governed by the usual mirror relation: angle of reflection equals angle of incidence. This reflection creates an image of the light which has the color of the light, not the color of the object. In the usual practical case, the light source is white. If the source is small, its image will be a small but bright highlight. If the source is large, its image will be large and dim, and it will be a "veiling reflection."

The 96 percent of the incident light that is not reflected at the surface bounces around among the pigment particles, so that part of it eventually scatters back out. This so-called body reflection differs from the surface reflection in two respects: it has a color determined by the pigment, and it is diffuse, rather than directional. The diffusely reflected light conveys the color that is characteristic of the object. If the surface contains pictures or print, those are also part of the body reflection. The highlight is always white, so it's not a source of color information about the object, though it probably is a source of information concerning object shape.

The rough object on the right in Figure 4 is less interesting, because the surface and body reflection are both diffuse. Highlights are absent, while there is a permanent veiling reflection that represents about four percent of the incident light, scattered in all directions. This veil of white light limits the extremes of black and deep color that rough surfaces can display.

Veiling Reflection Relative to White

Let us now look at the effect of luminaire size in an idealized case which expresses the essence of the veiling reflection problem. In order to do this thought-experiment, we need the thought-apparatus of Figure 5.

This apparatus consists of a flat, shiny black piece of glass next to a diffusely reflecting perfect white surface. Let us assume that there is one circular light source centered over the black-and-white display, and the observer is looking down on it. The observer's head is small, and things are misaligned enough that he doesn't cast a shadow on the white or see his own reflection in the black. The only luminance seen in the glass is the mirror-image of the source; this image has four percent the luminance of the source itself. By formulas that are in the handbooks, the luminance of the white surface can be calculated in terms of the source luminance and
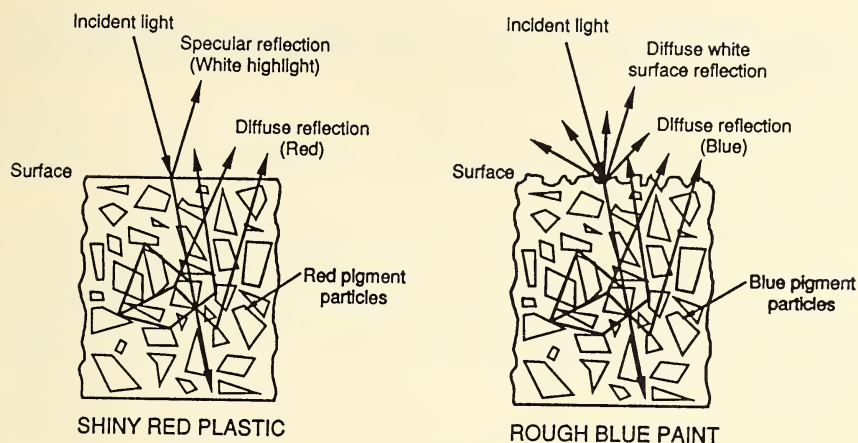
Figure 4. Classic textbook representations showing the role of surface reflections in light reflections from non-metallic objects. Whether the object is rough or smooth, approximately four percent of the incident ray is reflected at the surface with little color change. When the surface is smooth, this reflection is specular (mirror-like) and the eye can distinguish it from the diffuse but colored reflection by the pigment granules. When the surface is rough, the surface reflection is diffuse and merges with the "body" reflection from the pigment. These drawings are adapted from one in Richard S. Hunter's classic book, *the Measurement of Appearance*[4].
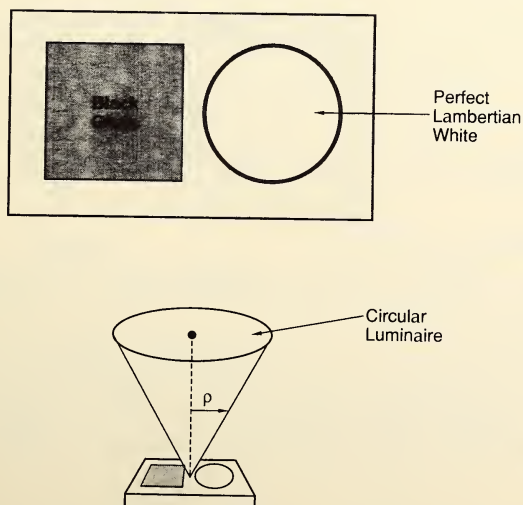
Figure 5. Idealized apparatus for "thought-experiments" concerning veiling reflections. A test object consists of a flat piece of black glass mounted next to a diffusely-reflecting ("Lambertian") white surface. In the basic thought-experiment, the test object is positioned under a circular luminaire, and the luminance of the luminaire's reflection in the black glass is compared to the luminance of the white.

the semi-subtense ρ. Then if we calculate the gray level of the black glass—its luminance divided by that of the white surface—we get a formula that's independent of source luminance:

$$g = f_g / \sin^2(\rho) \tag{2}$$

Here, g = "gray level" of the glass, ρ = semi-subtense of the circular luminaire, and f_g is the surface reflectance of the glass, which we of course take to be 0.04 (= 4%).

Equation (2) is graphed in Figure 6 as $\log_{10}(g)$ versus ρ, with g in percent. As ρ reaches 90°, meaning that the light source becomes a hemisphere, the gray level becomes four percent. This is the minimum veiling reflection amplitude due to a uniform source. As ρ becomes smaller and smaller, the image luminance increases without limit, meaning that highlights due to small light sources can have luminances much greater than the white.

Nonlinearity of Lightness Perception

We just saw that if the viewer does not avoid the veiling reflection—get it off the task—the minimum luminance it can have is four percent of white. It might appear that four percent is a negligible gray level, but this is not so. Four percent may be large in relation to the gray levels of details that are veiled. Also, the eye sees blacks, grays and whites in a non-linear way, such that the bottom few percent in gray level are of major importance to perception, while the top few percent are not.

Practical schemes for assigning numbers to object colors, such as the Munsell and CIELAB systems, account for this non-proportional response. It can be approximated by saying that perceived lightness is proportional to the cube root of gray level. In the Munsell system, the four percent veiling reflection corresponds to Munsell value 2.3, or 23 percent of the way from extreme black to perfect white.

The Three Four-Percent Rules

We have now stated three four-percent rules:

1. At a shiny air-dielectric interface, about four percent of an incident light beam will be reflected.

2. The luminance of a source image (veiling reflection or highlight) in a shiny dielectric is about four percent of the source luminance.

3. Under hemispherical lighting (or spherical, for that matter), the veiling reflection has a gray level of four percent, or a Munsell value of 2.3.

Veiling Reflections and the Color Gamut

We have computed veiling reflections as a fraction of a reference white. Because of its adaptive ability, the eye generally sees object colors in relation to a reference white—even when a reference white is not in the field of
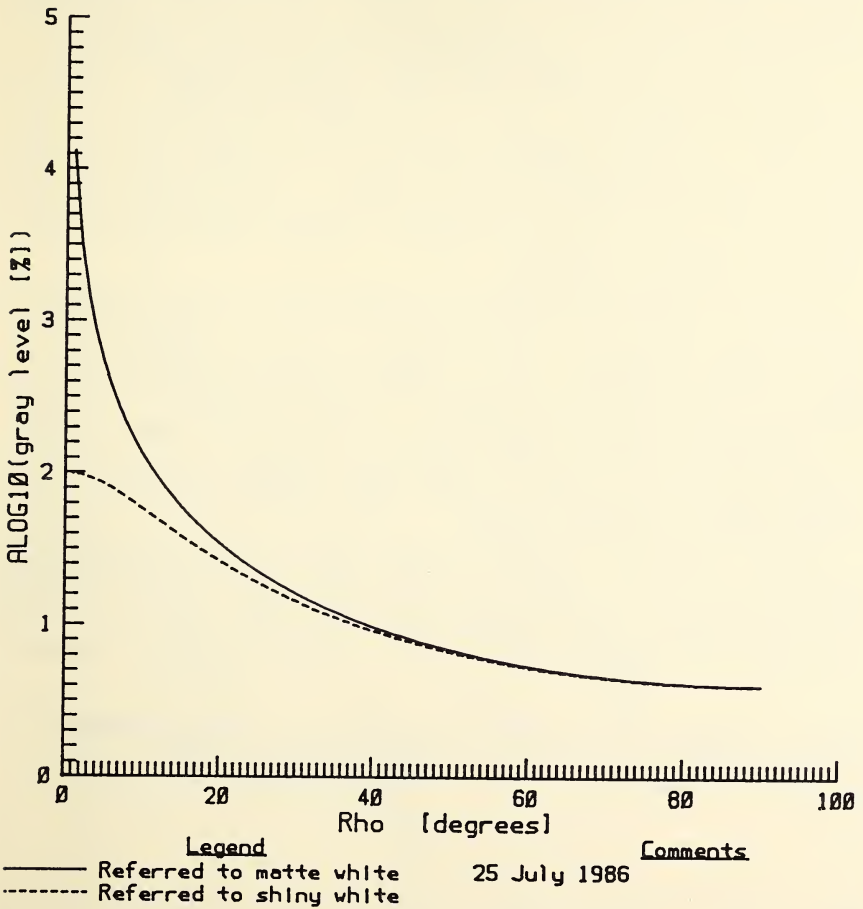
Figure 6. Logarithm to base 10 of the gray level of black glass, as a function of the luminaire size parameter ρ.

view.[5] The range of object colors that can be seen may be pictured as a "color solid," with white at the top, black at the bottom, and an asymmetric curved surface bounding the most saturated colors around the outside. Blacks, grays, and whites run from bottom to top up the middle of the solid, with colors becoming increasingly saturated toward the outside.

In other words, to make a color solid, you first establish a scheme to describe hue, saturation (= chroma) and lightness (= value). Then you work out the limits on saturation as a function of value for each hue. The color solid will display such basic features of color vision as the fact that the most saturated yellows are light, while the most saturated reds and blues are dark. An illustrated description of some color-naming schemes and their associated color solids appears with the entry "color" in the American Heritage Dictionary[6].

One way to determine the limits of saturation is simply to search for examples of extreme colors. Michael Pointer did this, using Illuminant C as the reference illuminant and CIELAB as his color-quantification scheme[7]. Using the cylindrical polar version of CIELAB, he reported his findings as sets of points in constant-hue planes, for every ten degrees of hue angle.

Starting with Pointer's points, I calculated the effect of veiling reflection on each limiting color point. The solid lines in Figure 7 connect Pointer's points. The chains of arrows show the effect of veiling reflections in successive increments of four percent.

Figure 7 shows only four constant-hue planes, but similar graphs were computed for each ten degrees in hue angle. The net effect of veiling reflections is to reduce the volume of the color solid. Table 2, calculated from the complete sets of computed points, shows how the volume of the color solid—the range of colors available—is reduced by veiling reflection.

Table 2, Volume of Color Solid as a Function of Veiling Luminance

Veiling Reflection, As Percent of White	Relative Volume of Color Solid, as Percent.
0 %	100%
4	63
8	46
12	35
16	28

From this table we see that a four percent veiling reflection, the veiling luminance expected under spherical lighting, causes a 37 percent reduction in the range of colors seen.

The Matte versus Glossy Issue, Again

It was stated earlier that a matte surface gives a fixed level of veiling reflections, comparable to that under spherical lighting, or four percent of white. From Figure 7, we see that this leads to a considerable loss of dark, saturated colors—-the colors that artists call "deep."
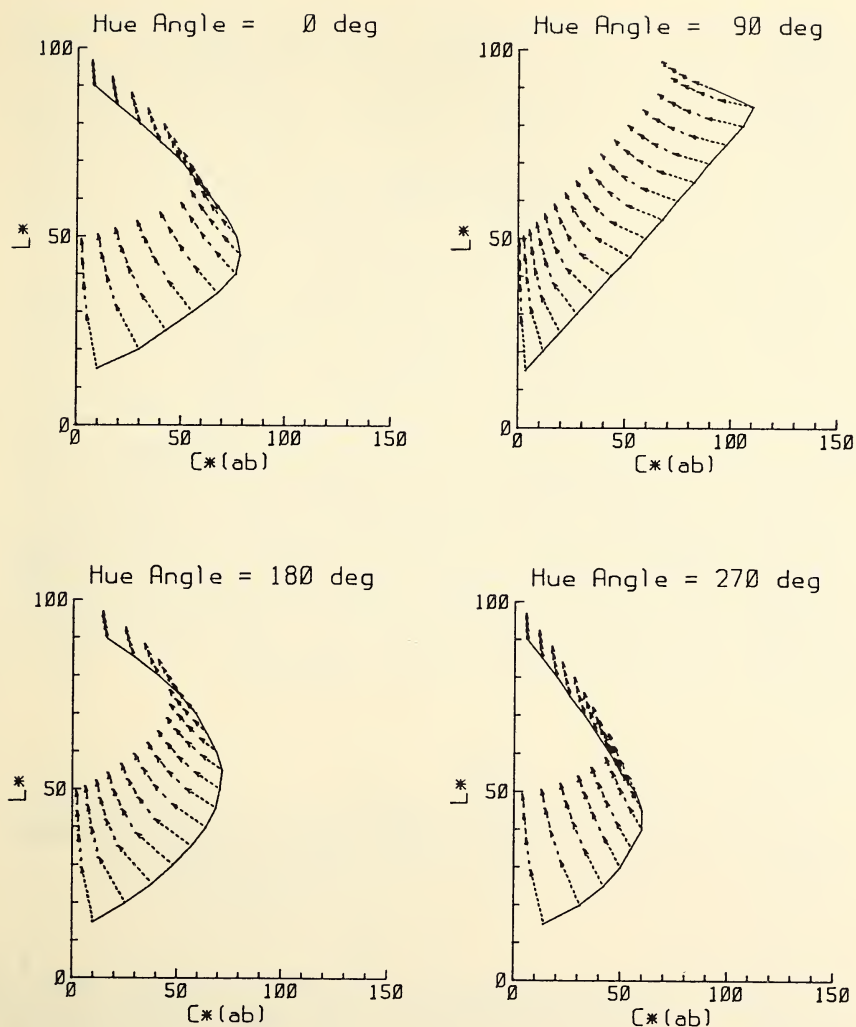
Figure 7. Reduction in the range of lightness and saturation of surface colors due to veiling reflections. Data are plotted in the cylindrical-polar version of the CIELAB uniform color space, at four selected hue angles (h^*). Radial coordinate c^* represents saturation, while axial coordinate L^* is a measure of lightness. The solid lines represent the limits attainable with real pigments, according to Pointer[7]. Chains of arrows show successive shifts as veiling reflection is increased to 4%, 8%, 12%, and 16% of white. At these selected hue angles, the shifts stay in the constant-hue plane, but at other h^* values, they don't. At all hue angles, including those not shown, a similar systematic loss of saturation and of lightness range occurs.

The decrease in possible saturation of colors can be seen in the differences between the matte and glossy sets of Munsell papers. Munsell papers are a commercially available[8] set of color-painted papers based on the Munsell uniform color space[9,10]. There are currently about 1600 chips in the glossy finish collection, but only 1300 in the matte finish collection. Because the published sets are not uniformly spaced in the three-dimensional color space, these numbers are not a direct measure of the volume of glossy and matte color solids, but do give some clue as to the practical loss of variety in matte colors. Since Munsell notation is inherently a cylindrical system, similar to the cylindrical version of CIELAB, the nature of the matte-glossy difference can be exemplified in diagrams at constant hue, similar to Figure 7. Figure 8 shows two constant-hue slices through the set of available papers, with available glossy papers denoted by rectangles. For instance, the rightmost rectangle in the top row under "10Y" indicates that there is a glossy paper for notation 10Y 9/6. (10Y indicates hue, a yellow; 9/ tells the Munsell value; and 6 indicates chroma or saturation.) The heavy line encloses those rectangles for which a matte paper is available. At hue 10Y, we see that eight hues are lost, but two are gained in going from glossy to matte. At hue 2.5R (a red), nine hues are lost and none are gained. The solid line bisects the rectangle for 2.5R 2/2 to indicate that a chip is sold at 2.5R 2.5/2. Where possible, matte chips are made with value 2.5, since none can be made with value 2. In the neutral series, the pure whites, grays and blacks, the glossy series goes down to value 0.5/, while the matte series stops at 1.75/ on the black end.

Poor Man's Antireflection Coating

We may conclude that a shiny surface is a poor man's anti-reflection coating. To use this anti-reflection coating, the poor man must have a dark surface near his light source that he can put at the specular angle.

Color Rendering

The loss of saturated colors due to veiling reflections adds to any loss that is due to color-rendering deficiencies of the light[11]. For instance, Xu[12] found that a Warm White fluorescent lamp reduces the volume of accessible colors by 25% from its value under Standard Illuminant A. Xu's concept of a color solid volume is similar to that in the present paper, although his calculation differs in important details.

Highlights and Luminance Dynamic Range

Figure 5 and Equation 2 above were introduced in order to quantify veiling reflections. However, the same equation applies to highlights— the images of compact sources. That is, highlight luminance, when ex-
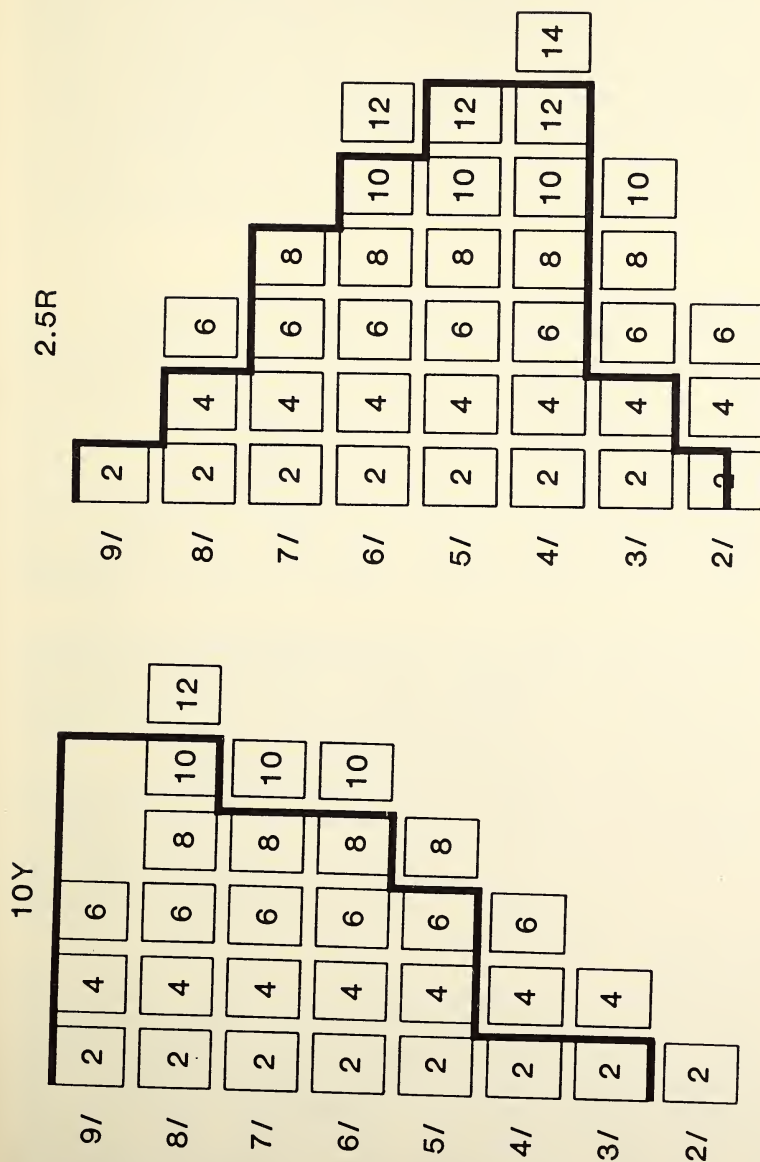
Figure 8. Comparison of the range of Munsell colors available in the glossy and matte sets. Two constant-hue planes (10Y and 2.5R) are shown, similar to Figure 7. Rectangles with numbers stand for glossy chips that are available. The heavy line indicates the range of matte chips available. For instance matte chips are available, but not glossy ones for notations 10Y9/8 and 10Y9/10. Glossy papers, but not matte, are available for 10Y4/4 and 10Y4/6.

pressed as a multiple of the luminance of a diffuse white surface, will get smaller as the source area increases.

Whereas veiling reflections reduce the range of luminances in a scene by washing out blacks and deep colors, highlights add to dynamic range. Table 3 gives calculated highlight luminance and dynamic range for the same six lights listed in Table 1. "Dynamic range" is used here to mean the ratio of the highest luminance to the lowest in a scene. With the black-glass-plus-white diffuser test object of Figure 5, both the high and low luminances would occur in the black glass, except under the luminous ceiling. In that case, an unavoidable veiling reflection would make the black glass less black, while the brightest area would be the white surface.

Table 3: Highlight Luminance and Dynamic Range

Light Source	Highlight Luminance	Dynamic Range
Unfrosted 60 W	25,000	2,500,000
Sun	1900	190,000
Ordinary frosted 60 W	1600	160,000
Soft White	210	21,000
Single fluorescent tube	11	1100
Luminous Ceiling	.05	20

I have assumed the black glass without veiling reflection to have a luminance of 0.01; there's a slight haze on the surface. This is inconsistent with previous graphs, but is a realistic way to avoid saying that dynamic range becomes infinite. I have assumed that a dark surface was available to put at the mirror angle, except under the luminous ceiling.

The assumption of a test object lit by a single unfrosted bulb is a bit artificial, though it indicates what is possible with, say, compact spotlights. Putting aside the unfrosted bulb, consider the transition from daylight to a luminous ceiling. The sun's reflection in the black glass has 1900 times the luminance of the diffuse white. The luminous ceiling's reflection is only 0.04 of white, giving a dynamic range which is less than that in daylight by a factor of 10,000.

These calculations concerning amplitudes of veiling reflections and highlights, and their effect on dynamic range can be reduced to man-in-the-street terms. Suppose that the man in the street is standing in front of a drugstore, diffusely lit by fluorescent lamps, looking in the plate glass windows. Suppose that it's a clear day, but the sun is fairly low in the western sky, so that the mean luminance of the outdoor scene is equal to that inside the drugstore. What the man in the street will see, or what you and I will see is that the scene in the drugstore looks "washed out," compared to the scene outdoors.

What we now understand from calculation is that the drugstore does not look washed out for some mysterious reason involving fluorescence of the crystalline lens, or flicker, or some quirk of the visual system. It looks washed out because it *is* washed out. Highlights are dim and large; blacks and saturated colors are covered by veiling reflections. This is in addition to the loss of color contrast because of the color rendering inferiority of fluorescent lights, the loss of black-white contrast because of the lack of shadows in the drugstore, and the enhancement of color contrast outdoors due to the fact that light from the west is reddish while that from the east is bluish.

Summary

The effects of light source size on the contrasts of illuminated objects have been examined by methods of basic optics and visual science. It was observed that practical light sources vary by a factor of 100,000 or more in the solid angle they subtend (Table 1). Equation (1) shows how far a book must be tipped to avoid veiling reflections, as a function of luminaire size. The discussion of Figure 4 reviewed the classic picture of light reflection from non-metallic objects. The "surface reflection" from a shiny object is directional and white, while the "body reflection" is diffuse and has the characteristic color of the object. The rough object shows surface reflection also, but it is diffuse and therefore becomes merged with the colored body reflection.

Figure 5 presented an apparatus for "thought experiments" regarding the amplitude of veiling reflections. The apparatus has a black glass next to a white diffuser. Equation (2) gives veiling luminance, expressed in units of the white object's luminance, as a function of a circular luminaire's semi-subtense. That is, the absolute luminance of the luminaire drops out and luminaire size is the important independent variable. The veiling reflection due to the uniformly luminous source reaches a minimum when the source is a hemisphere—equivalent to a complete sphere since the objects are flat. Equation (2) also applies to highlights, and it shows that the source image (highlight) in the black glass can have a much higher luminance than white, as the source becomes smaller and smaller.

Three "four percent rules" were found, which give a simple way to estimate the amplitude of veiling reflections. The dimmest image of a uniform source occurs when the source is the biggest, namely a hemispherical or spherical source. The third four-percent rule says that this minimum veiling luminance is four percent of white. While this percentage may sound small, it is often not small in comparison to the details that it veils. Also, because of the nonlinearity of human lightness perception, a four percent gray level is perceptually 23 percent of the way from extreme black to pure white.

Calculating the effect of veiling reflections on the gamut of colors that can be seen showed that just the four percent veiling luminance eliminates

many deep colors, for a net loss of 37 percent of the range of color perception (Table 2). Deep colors are lost under any lighting when objects are given a matte surface. The everyday commercial reality of this observation was shown by Figure 8, a comparison of some pages in the matte and glossy versions of the Munsell book. The subject of color rendering was raised, only to say that losses of color contrast due to color rendering deficits occur in addition to the loss by veiling reflections.

The transition from sunlight to a luminous ceiling turns highlights into veiling reflections, compressing the range of luminances from both ends. Table 3, based on simple calculations, shows that the dynamic range is reduced by a factor of about 10,000. This finally led to the statement that when the whole ceiling is bright with sources of mediocre color rendering, the scene will look washed out because it is washed out.

Conclusion

The reader may well ask, "Is all this really new? Has no one noticed before that filaments are tiny and bright, while fluorescent tubes are four feet long and an inch and a half in diameter?" Of course, engineers have noticed these facts. When lamp manufacturers design a spotlight, they are well aware that what they are selling is the compactness of the source. Also, in my analysis, I have adhered to traditional textbook visual science, indeed to such workaday vision results as the Munsell system; there's nothing new there. In fact, nearly all the background information for this chapter can be found in Wyszecki and Stiles's book[9] (the more recent edition would be better).

The only novel element has been to develop a theory of lighting quality based directly on the observation that source size varies widely. To my knowledge, existing textbooks fail to do this.

Acknowledgement

encouragement that we received early in our careers from Gordon Heath and the other dedicated faculty of IU Optometry.

References

1. Worthy, James A., Geometry and amplitude of veiling reflections, J. Illum. Eng. Soc. of North America, in press.
2. Jackson, John David, Classical Electrodynamics (John Wiley, New York, 1965).
3. Jenkins, F. A. and White, H. E., Fundamentals of Optics, Third Edition (McGraw-Hill, New York, 1957).
4. Hunter, Richard S., The Measurement of Appearance (Wiley-Interscience, New York, 1975).
5. Worthey, James A., Limitations of color constancy, J. Opt. Soc. Am. **A 2**, 1014-1026 (1985).
6. Morris, W., Editor, The American Heritage Dictionary of the English Language (Houghton Mifflin, Boston, 1978).
7. Pointer, Michael R., The gamut of real surface colors, Color Res. Appl. **5**, 145-155 (1980).
8. Munsell Color Company, 2441 North Calvert Street, Baltimore, MD 21218, U.S.A.
9. Wyszecki, Gunter and Stiles, W. S., Color Science: Concepts and Methods, Quantitative Data and Formulas, (John Wiley, New York, 1967).
10. Newhall, S. M., Nickerson D. and Judd, D. B., Final report of the O.S.A. subcommittee on spacing of the Munsell colors, J. Opt. Soc. Am. **33**, 385-418 (1943).
11. Worthey, James A., Opponent-colors approach to color rendering, J. Opt. Soc. Am. **72**, 74-82 (1982).
12. Xu, H., Color-rendering capacity of illumination, J. Opt. Soc. Am. **73**, 1709-1713 (1983).

SHOULD EXPOSURE TO SOLAR UVR BE A CONCERN?
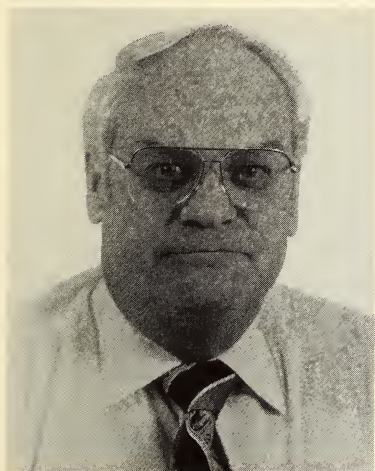
Donald G. Pitts
Professor of Environmental Optometry and Visual Science at University of Houston College of Optometry, Houston, Texas. He is consultant to NASA on space vision and has served on the National Academy of Sciences/National Research Council Committee on Vision.

ABSTRACT.   Descriptions of solar spectral radiation and the changes it undergoes during passage to earth are presented. With the projected decreases in the ozone layer, changes are found in UVR below 340 nm that include an increase in the level of UV irradiance and the passage of lower wavelengths of UVR to earth. Research on the action spectra for the cornea, lens and retina from exposure to UV radiation is reviewed. The data demonstrate that ocular damage occurs at UV exposure levels very close to that contained in sunlight and that exposure of the eye to solar radiation is hazardous. Methodology for protecting the eye against UV exposure by using the appropriate ophthalmic lenses, contact lenses or intraocular lenses is described. A list of activities that require the patient to be protected from UVR exposure is provided.

SHOULD OCULAR EXPOSURE TO SOLAR UVR BE A CONCERN?

Introduction

The fact that ultraviolet radiation (UVR) exposure to the eye results in damage has been known for centuries since Xenophon used the term "snowblindness" in his treatise *Anabasis* ca. 370 BC. Later, the accounts of the march of Hannibal over the Alps in 218 BC during the Second Punic War against Italy clearly describe the classical signs and symptoms of photokeratitis in the soldiers, the horses, and the elephants. Since these ancient descriptions in history, ultraviolet radiation and its effects on the human eye have been studied sporadically.

Widmark[1,2,3] was the first to accomplish research on the effects of exposure of UVR to the cornea and, in fact, named the hazard as photo- or UV keratitis. In addition to Widmark, a number of researchers detailed the corneal epithelial response describing cellular degeneration including the nucleus, inhibition of epithelial mitosis, eosinophilic staining and cell death (Verhoeff and Bell[4]; Duke-Elder and Duke-Elder[5]; Buschke et al[6], and Friedenwald et al[7]). Cogan and Kinsey[8] provided the most reliable quantified data from their laboratory studies of UVR exposure to the cornea with their corneal threshold value of 0.015 J cm^{-2} at 288 nm for the pigmented rabbit eye.

The present day research efforts on UV effects began with the Apollo program in 1964 when extra-vehicular activity became a necessity for exploration of the surface of the moon. The resulting technology has led not only to protective devices but also to additional avenues of effort. For example, the present emphasis on ozone research has been aided by measurements from space satellites. The awareness of the cholorofluorocarbons (CFCs) as the cause of the atmospheric decrease in ozone has resulted from the awareness of the dangers from an increase in ultraviolet radiation at the earth's surface.

It is the purpose of this paper to review briefly the action spectra for UV exposure to the eye, describe the spectral solar irradiance from the sun, look at the effect of increased and decreased ozone levels on solar UVR, to address criteria to avoid exposure to the human eye, and illustrate

SOLAR RADIATION
SOLAR CONSTANT: 1367 W/M^2
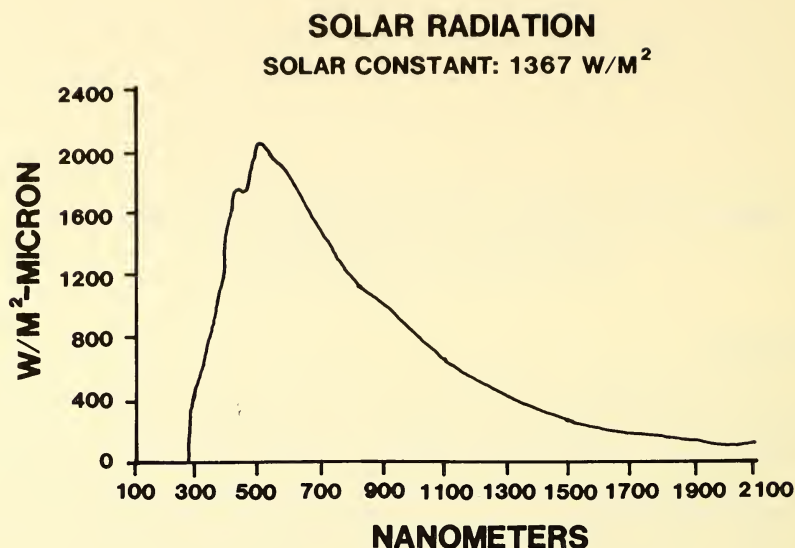
Figure 1. The solar spectral irradiance for the solar constant of 1367 W/m². (After Mecherikunnel et al[14]; Neckel and Labs[13]).

present technology available to protect the eye. The data presented represent research that was initiated in 1964 and has continued to the present day.

Sunlight as a UV Source

A human on the earth's surface is exposed to UVR that is highly directional from the solar disc, both diffuse and directional from the sky and by diffuse and specular reflection from the various surfaces of the earth. Each of these factors continuously varies with the time of day, zenith angle of the sun, the cloud cover and changes in reflecting surfaces. These variables make the determination of solar UVR available for exposure very difficult (Rosenthal, et al[9,10]; and Scotto[11]). Based on albedo of the surfaces, the total radiation reflected from grass is about 3%, 20% to 30% from sand and 3% to 5% from fresh water, but reaches 85% to 95% from fresh snow.

If the total solar radiant energy were measured in space, the value would vary from 1353 Wm⁻² to 1369 Wm⁻². The solar radiation received on earth from the solar disc has been called "direct" while the solar radiation scattered as the result of water vapor, turbidity and Rayleigh scatter is called "global" or "indirect" solar radiation. In the ultraviolet region of the
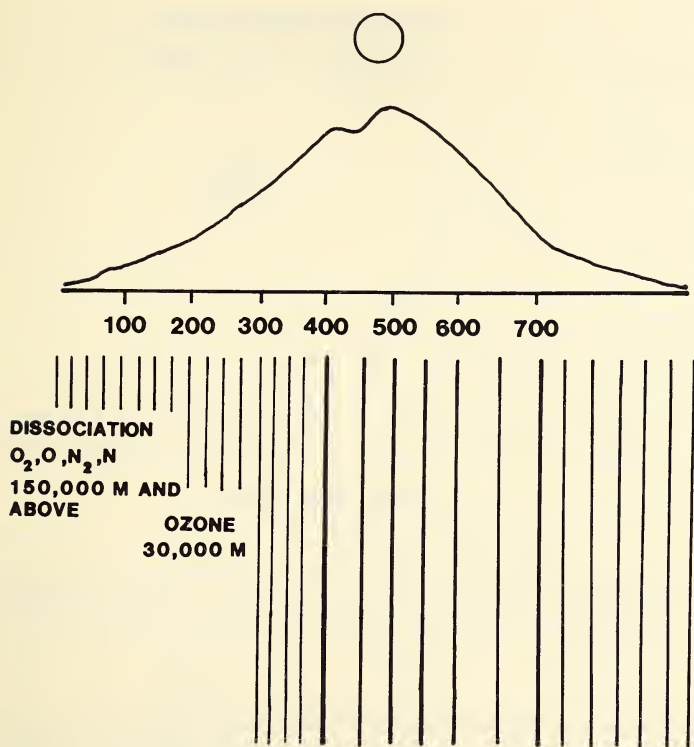
Figure 2. The filtering effects of atmosphere on solar radiation prior to it reaching earth and being called sunlight. The text explains the filtering by oxygen, nitrogen and ozone.

spectrum, the global or indirect radiation exceeds that from the solar disc. The total spectral radiation includes both the direct solar radiation and the global solar radiation. The total solar power measured in space is called the solar constant (Frohlich and Brusa[12]).

The solar spectral irradiance is the distribution of the solar constant as a function of wavelength. (Neckel and Labs[13]; Mecherikunnel et al[14]). Figure 1 presents the solar spectral irradiance for a solar constant of 1367 Wm^{-2}. Over 96% of the solar spectral irradiance is contained in the wavelength range beginning at 270nm and extending to 2600nm while 49.6% lies in the visible spectrum in the 400nm to 760nm waveband. It is remarkable that half of the solar radiation is in the wavelength range that serves the human for vision.

As solar radiaiton approaches earth, its spectral characteristics undergo rather startling changes. Absorption in the spectral region below 85nm is

Donald G. Pitts

SOLAR RADIATION
ON EARTH: 726.9 W/M^2
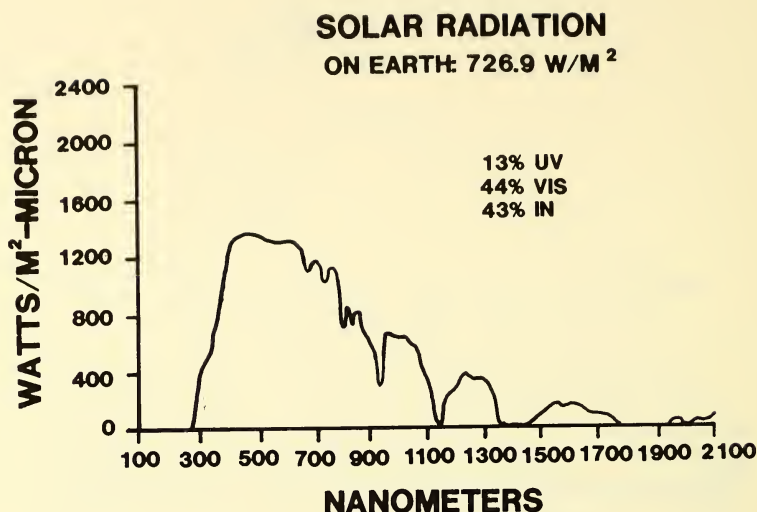
Figure 3. The spectral irradiance of sun on earth. The measurements represent sunlight after it has been filtered by the atmosphere and is incident on the earth. The effect of the filtering was to limit the short wavelength to 288nm and to produce absorption bands in the infrared.

due chiefly to atomic oxygen (0), molecular oxygen (0_2), atomic nitrogen (N) and molecular nitrogen (N_2). The absorption between 100,000 km to 30,000 km for the 85nm to 200nm waveband is principally by molecular oxygen 0_2. At 30,000 km and below, ozone (0_3) absorbs the remaining solar UVR in the 200nm to 288nm wavelength range (Figure 2). Thus, essentially all of the solar UV radiation above 288nm is absorbed by the earth's atmosphere and its spectral distribution is dramatically different from that in space (Figure 3). It becomes rather simple to understand that small changes in oxygen, nitrogen and ozone may result in a relatively small increase in the total solar radiation but a disporportionate increase in the UV-B radiation that would increase the hazards of UVR on earth rather dramatically.

Unlike solar UV, solar infrared radiation does not interact with ozone, oxygen or nitrogen on its propagation to earth. However, infrared is ob-sorbed by the atmospheric moisture, dust, $C0_2$ and other impurities and its total energy is greatly reduced before it reaches earth. These interactions result in the absorption bands seen in the infrared wavelengths of the solar spectral radiation. Thus, the solar radiation on earth that we call sunlight consists of approximately 13% UVR, 44% VIS and 43% IR.

The above descriptions should explain why it is easy to obtain a severe sunburn on a cloudy day. Infrared radiation is the portion of the solar

spectrum that causes the skin to "feel" hot during sunbathing. It is the ultraviolet radiation that is responsible for the photochemical reactions to exposure that we call suntan and sunburn; however, UVR does not make the skin "feel" hot because the response of the skin to UVR exposure is delayed 6 to 12 hours. On a cloudy or overcast day, a greater portion of infrared is absorbed by the moisture in the clouds, consequently, the sun does not feel "hot" while the UVR is readily transmitted and allowed to do its harm. Since the normal warning signs of IR are absent, the results are an unusually severe sunburn.

Shettle and Green[15,16] have published data on global and direct ultraviolet radiation in the 280nm to 340nm wavelength range. These data were measured for a horizontal surface on earth for different degrees of solar elevation. Table 1 presents part of their data for a solar elevation of 90° or with the sun at zenith and for three different thicknesses of the ozone layer.

WAVELENGTH IN NANOMETERS	SOLAR ELEVATION - 90 DEGREES		
	OZONE THICKNESS		
	0.24 atm-cm	0.32 atm-cm	0.40 atm-cm
	IRRADIANCE IN W m^{-2} nm^{-1}		
280	1.54×10^{-14}	9.09×10^{-18}	5.44×10^{-23}
285	1.78×10^{-8}	9.50×10^{-11}	5.12×10^{-13}
290	3.58×10^{-5}	2.14×10^{-6}	1.29×10^{-7}
295	2.36×10^{-3}	5.14×10^{-4}	1.13×10^{-4}
300	2.45×10^{-2}	1.07×10^{-2}	4.72×10^{-3}
305	9.25×10^{-2}	5.91×10^{-2}	3.79×10^{-2}
310	2.02×10^{-1}	1.58×10^{-1}	1.24×10^{-1}
315	3.24×10^{-1}	2.48×10^{-1}	2.49×10^{-1}
320	4.39×10^{-1}	4.09×10^{-1}	3.81×10^{-1}
325	5.41×10^{-4}	5.20×10^{-1}	5.01×10^{-1}
330	6.30×10^{-1}	6.17×10^{-1}	6.05×10^{-1}
340	7.87×10^{-1}	7.82×10^{-1}	7.77×10^{-1}
290-320	5.41 Wm^2	4.61 Wm^2	3.99 Wm^2
280-340	15.2 Wm^2	14.2 Wm^2	13.4 Wm^2

TABLE 1. Global solar UV radiation reaching a horizontal surface on the ground for the solar elevation of 90 degrees with ozone thicknesses of 0.24 atm-cm, 0.32 atm-cm and 0.40 atm-cm. The solar elevation of 90 degrees is equivalent to a zenith angle of 0 degrees. The standard atmospheric ozone thickness is taken to be 0.32 atm-cm. (Data from Shettle and Green)

The Ozone Story

The development of the supersonic transport in the USA (SST) created a concern that the pollutants from their jet engines could result in a degradation of the environment in the stratosphere. The result has been a

series of studies and reports beginning in 1974 that were designed to ana-
lyze the impacts of the propulsion effluents of aircraft engines on the
stratosphere and the impacts of the climatic changes on the biologic systems
on earth (Anon[17]). These studies concluded that such climatic changes
would result in an increase in solar UV radiation due to partial loss of
ozone and changes in temperature and precipitation associated with the
increase in pollutants—aerosol sulphates and water vapor.

The situation is not as clear as the introductory paragraph appears to
make it. In 1974, Molina and Rowland[18] suggested that the chlorofluro-
carbons, (CFCs) also called chlorofluoromethanes (CFMs) could potentially
harm the stratosphere by interacting with ozone to produce oxygen; thereby,
reducing the stratospheric ozone. The result would be a net increase in
the ultraviolet radiation reaching earth with deleterious effects on plants,
animals and humans. The irony was that a mathematical model (Halpern
and Braslau[19]) based on a fixed amount of ozone had predicted a significant
decrease in UVR and the visible spectrum on the earth.

Using the effluent production of the nitrogen oxides (Nox) by the SST,
Alyea et al[20] predicted an annual depletion of 12% in the stratospheric
ozone in the northern hemisphere and a depletion of 8% in the southern
hemisphere. By 1979, Rowland and Molino's hypothesis had become ac-
cepted and different studies had established ozone depletion rates of 7.5%,
10.8% and 16.5% if the release of CFC's were continued at the 1977 rate
(Anon[17]). More recently Gille et al[21] have reported a 12% reduction in
stratospheric ozone over an 11 year solar cycle or a 1.09% reduction per
year.

Kerr[22,23] reviewed the ozone problem and emphasized several impor-
tant points. The intensive program to re-evaluate the ozone measurements
shows rather conclusively that the stratospheric ozone level has decreased
over the past 17 years. An unexpected result was that there were larger
than expected decreases in ozone in the colder climates and higher lati-
tudes. This finding suggests that the process may be mediated by the ice-
chemistry phenomenon. The Antarctic hole reported a loss of 50% of the
total ozone in the hole and a 95% loss in the lower stratosphere in October
1987. In addition, there were record losses of ozone outside the hole, the
hole lasted longer, the hole was colder than other years, and there has
been a total decrease in ozone at latitude 60's and further south the entire
year. The weight of evidence is that the CFC's are primarily responsible
for these observed changes in the stratospheric ozone.

The importance of the ozone story to us is that the lower wavelength
and the intensity of the UV irradiance falling on the earth depends on the
thickness of the ozone layer and the air mass of the atmosphere. Shettle
and Green[15,16] calculated the global, direct and diffuse solar UV irradiance
reaching the ground. They incremented the ozone layer thickness by 25%

steps beginning at 0.16 atm-cm and extending to 0.40 atm-cm. Their global irradiance data for solar elevations of 40°, 60° and 90° and ozone layer thicknesses 0.24 atm-cm, 0.32 atm-cm and 0.40 atm-cm are plotted in Figure 4. These data provide an insight into the spectral irradiance changes brought on by the changes in the ozone layer thickness. The curves asymptote at about 340nm indicating that the changes brought about by changes in the dozone layer occur below 340nm. As the ozone layer reduces its thickness equivalency, there is a relatively larger increase in the short wavelength ultraviolet. Finally, as the ozone layer decreases, the lowest UV wavelength that reaches earth shifts from 296 nm to about 288 nm. Thus, there is a significant increase in the UVR waveband, the level of irradiance and, at the same time, a shift to higher energy photons in the lower wavelength UVR.

The Action Spectrum of the Eye:

Since this presentation has limited time, it will be presumed that the transmittance of the ocular media is well known to the audience and, as a consequence, transmittance data will not be presented. However, there are certain terms which need to be understood for the remainder of the paper to be followed coherently. The first term is the watt (W) which is the unit of power. If a watt is consumed in one second, a Joule (J) of energy has been expended. When biological systems are exposed to radiation, the number of watts (W) consumed during the exposure duration (s) results in a radiant exposure $H = W \cdot s$.

The ultraviolet spectrum has been sectioned arbitrarily into three divisions by the CIE: UV-C from 200nm to 290nm; UV-B from 290nm to 320nm; and, UV-A from 320nm to 380nm. There were no compelling reasons for such division but it applies nicely to ocular damage to the different portions of the eye as UV exposure will be discussed. The quantum energy of a photon is inversely proportional to the wavelength, with the electron volts (eV) of the photon varying from 6.2eV to 3.1eV as we pass from 200nm to 400nm. Absorptions of photons with energy levels from 1.0 to 4.0 eV result in photochemical reactions. At 3.5eV the carbon-nitrogen bond is broken, at 4.3eV there is disruption of the carbon-hydrogen bond and the carbon-carbon double bond can be broken by eV. The absorption of radiation by a molecule is determined by its structure and the absorption or capture of a single photon usually results in a photochemical reaction. With the preliminaries completed, we shall return to the action spectrum.

The action spectrum for exposure of the eye to UVR exposure is shown in Figure 5. The original data of Cogan and Kinsey[8] are shown by the open, upright triangles. The ordinate presents the threshold radiant exposure in $Jm^{-2} \times 10^4$ and the abscissa the wavelength range from 200 to 400 in nanometers (nm). The X's provide the human corneal threshold for photo-

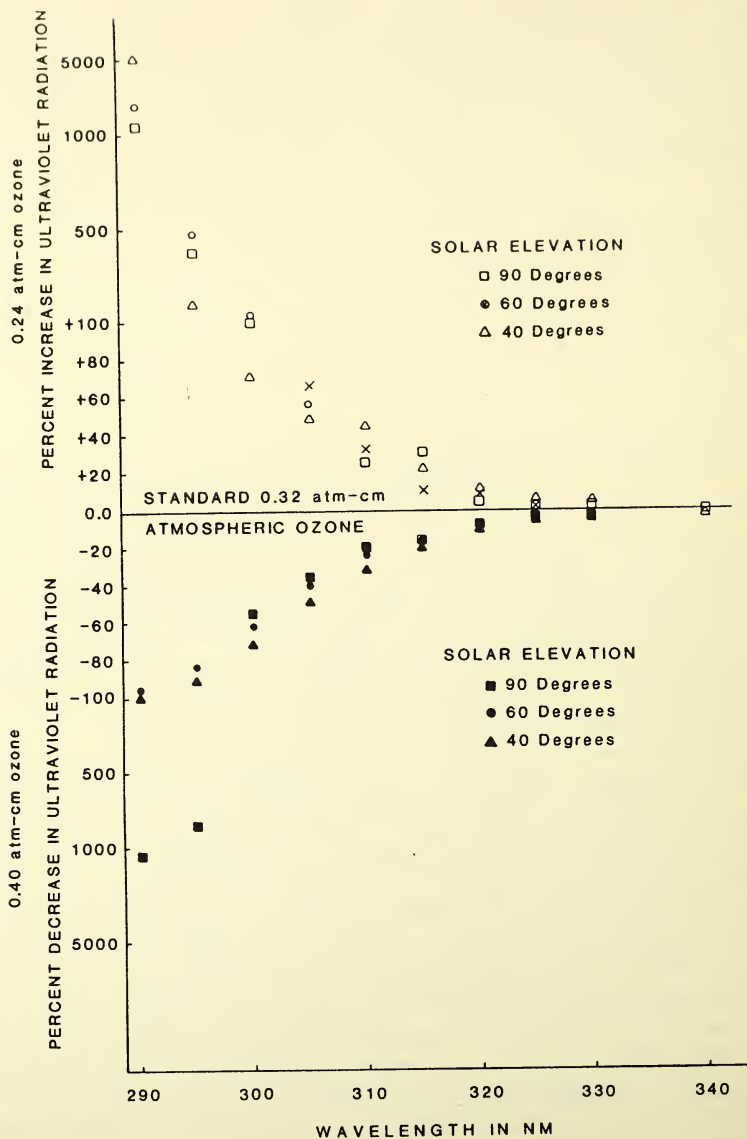**Figure 4.** The percent change in UVR resulting from ±25% change in the standard 0.32 atm-cm of ozone. The graph indicates that there are no UV changes above 340nm, large increases in UVR at wavelengths below 333nm and the wavelength limits of UVR reaching the are below 288nm. Each of these factors result in an increase in the abiotic wavelengths of UVR reaching the surface of the earth.

keratitis and its action spectrum covers the 210nm to 320nm wavelength range. The upper wavelength range for the human was terminated because of the lack of knowledge concerning UV induced cortical cataracts. The open, inverted triangles represent threshold data for the primate. Note the break at 320nm which is continued by open circles to 400nm. The open circle primate data were published by Kurtin and Zuclich[24] and, since they take the same form and values as the rabbit data, it is felt that the data for the three species of the human, primate and rabbit are not materially different above 310nm. If a value 400 times minimum threshold were used as the cutoff, the corneal action spectrum begins at 210nm and extends to 320nm. Obviously, corneal damage can be induced with UVR above 320nm but the threshold exposure levels are quite high; i.e., much higher than can be found in man made and natural radiation sources if the laser were excluded.

The squares represent the radiant exposure made to establish the UV induced cataractogenesis in the pigmented rabbit. The open squares provide exposures in which cataracts could not be induced while the closed squares represent the wavebands that produced acute cataracts. The wavelength range for UV induced cataractogenesis begins at 295nm and extends to 320nm. No cataracts were noted in either the pigmented rabbit or the primate for the wavelengths above 320nm.

The UVR data for retinal exposures are shown by the solid inverted triangles. The data of Ham et al[25] start at 320nm and extend to 400nm. The single retinal datum point at 300 nm is rabbit data from our laboratory but a complete spectral wavelength experiment has not been done.

Effects of Exposure of the Eye to Ultraviolet Radiation:

The action spectrum has been presented in the preceeding section. It is important to realize that the threshold values that were reported were based on a minimal damage criterion rather than a no-damage criterion. What this means is that for each threshold value there was found minimal ocular damage from the acute exposure that would demonstrate full recovery at some later date.

In the following paragrphs, an attempt will be made to describe research on the effects of exposure of the eye to UV radiation beginning with the cornea and proceeding to the retina. Data for the iris are almost totally lacking but in our experiments we have found swollen, sluggish iris, posterior synechia, loss of iridic pigment and hypertrophy of the iris but have not related the observations to an action spectrum. A preliminary histological study has affirmed clinical observations. Some day, we might take a more detailed look at the iris using the electron microscope or, hopefully, encourage a colleague to do likewise. Therefore, at the present time my discussion will be limited to the cornea, the lens and the retina.
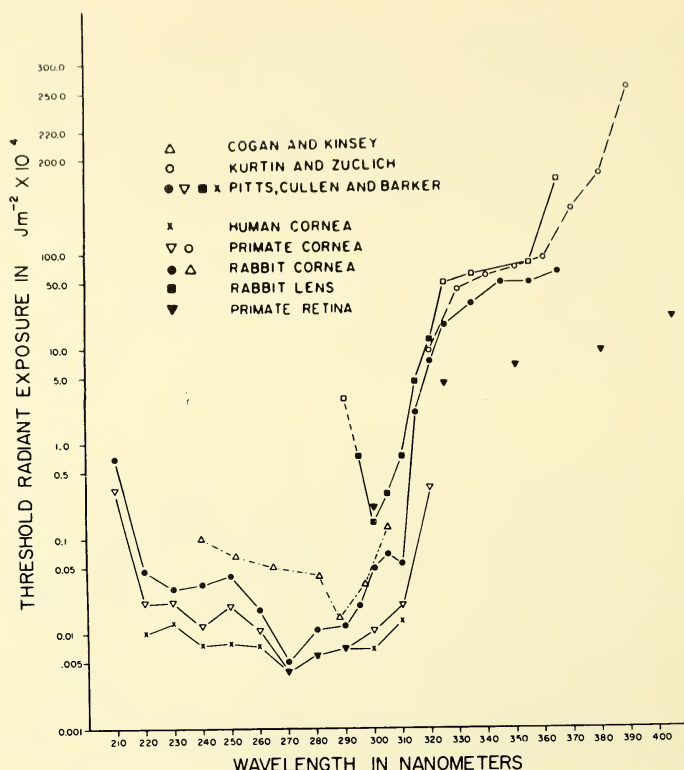
Figure 5. The action spectrum for exposure of the eye to ultraviolet radiation. The ordinate presents the threshold radiant exposure in J m^{-2} x 10^4 and is in scaled in log units. The abscissa gives the wavelength in nanometers beginning at 200nm and terminating at 400nm. The data symbols represent human, primate and rabbit exposures and include exposures to the cornea, retina and lens. The sources of the data are listed in the references.

Cornea:

Photokeratitis has been recognized as a primary response of the cornea to UVR since Widmark[1,2,3]. The action spectrum necessary to produce photokeraititis was defined by Pitts et al[26] and since then, corneal exposures have been shown to be oxygen dependent (Kurtin and Zuclich[27]), cumulative (Cullen[28]; Zuclich[29]) and to completely obey the law of reciprocity for both single exposures and multiple exposures (Zuclich and Connolly[30]). Pitts et al[31] studied the ability of the cornea to recover 8 days after exposure and found abnormal epithelial cells and, in addition, damaged and dead keratocytes were evident in the stroma. The endothelium demonstrated a reduced number of mitochondria, endoplasmic reticulum and occasional

vacuoles but normal corneal hydration was evident. They suggested that endothelial polymegethism occurs in the primate from UV exposure which has subsequently been confirmed in the human by Good and Schoessler.

Cullen, et al[32,33] and Doughty and Cullen[34] related UV exposure to endothelial damage using increased corneal thickness as an index. Doughty and Cullen[34] also showed that the damaged endothelium did not regain normal function even 3 months after exposure. Lattimore[35] found decreases in oxygen uptake, increases in glucose and glycogen but a decrease in phosphocreatine from UVR exposures to the corneal epithelium. The physiological responses were found within 2 minutes after exposure, were highly wavelength dependent and demonstrated a decrease in corneal epithelial metabolic activity. Blumthaler, et al[36] have reconstructed the UV exposures of skiers suffering from photokeratitis and demonstrated that laboratory data on animals can be used to predict human UV exposure effects. The levels of UV exposure used in these studies can be experienced from sunlight!

Crystalline Lens:

The data that define the action spectrum from 295nm to 320nm for the production of UV induced cataracts in the lens were presented by Pitts and Cullen[37]. Arguments that question the application of the data to the human have been raised since the results were derived with animals exposed to the UVR beam when the eye was directed toward the beam. This dilemma may be resolved by the epidemiological literature which indicates that the UV cataract is related to the exposure to sunlight (Hiller, et al[38]; Kahn et al[39]; Taylor[40]; Hollows and Moran[41]; Hiller, et al[42]; Leske and Sperduto[43]; and Brilliant et al[44]). The geographic areas with greater number of sunlight hours demonstrated a higher prevalence for cataracts than areas with fewer hours of sunlight. Although the issue is not entirely settled, the biochemical and epidemiological evidence combined with experimental data in animals make a causal relationship highly possible.

Considerable research on the biochemistry of the lens from UVR exposure has been done but will only be summarized briefly here. The literature (Pitts, et al[45]) indicates that biochemical alterations from UV exposure that lead to senile catacracts may occur via three major mechanisms: (1) Photo-oxidative damage to the lens crystallins which has been suggested to result in cross-linking, insolubilization and browning of proteins, especially in the lens nucleus. (2) Photo-oxidation of lens membrane lipids which may result in the inactivation of membrane functions, such as ATPase activity. In addition, photo-oxidation of membrane lipids has been suggested to initiate crosslinking and insolubilization of lens proteins. (3) Damage to the epithelium of the crystaline lens DNA has been proposed to result in the alteration of cell function and, also production of anomalous

cortical fiber proteins. Whether any or all of these mechanisms are, in fact, involved in cataract formation are yet to be established. Curiously, the final story may show that each of the proposed biochemical mechanisms plays its unique and, at the same time, combined role in the UV-induced cataract.

Retina:

Several facts are known about retinal exposure of UVR which point toward damage to the retina: UVR radiation near 320nm reaches the phakic retina, UV-A and UV-B impinge on the retina of the aphakic and pseudophakic eye (Kamel and Parker[46]; Saraux et al[47]) and the aphakic eye can see down to 365nm (Goodeve et al[48]). Schmidt and Zuchlich[49] reported a threshold radiant exposure of 0.36 J/cm^2 at 325nm measureed at the plane of the cornea but most of the data on UV retinal lesions show 5 J/cm^2 or higher retinal threshold level (Ham et al[25]). Using the absolute visual thresholds as a measure, Henton and Sykes[50, 51] reported a 3 to 5 times increase in the threshold immediately after exposure to 0.381 J cm^{-2}, 350nm, UVR that stabilized at a 0.5 log unit elevation after 3 to 5 days. They also found a 15 to 20% reduction in the number of receptor outer segments and nuclei. Rapp, Jose and Pitts[52] reported that ^{3}H-Thymidine was incorporated into the pigmented rat retinal outer nuclear layer, inner nuclear layer and ganglion cell layer which were taken as evidence that 300nm UVR was transmitted to and damaged the retina.

The aphakic eye is particularly suspect in damage from UV-B and UV-A exposures because the crystalline lens absorbed most of these wavelengths. After cataract surgery, chromatopsia and retinal blanching are found clinically and represent a photo-retinitis which could be prevented by protective ophthalmic lenses, contact lenses or intraculear lenses. Thus, it is obvious that UVR does impinge on the retina, affects the receptors, and is quite hazardous to the retina. However, more research needs to be done to establish the definitive action spectrum so that more adequate protection criteria for the retina may be developed.

Protection Against UVR

Protection against UVR may be accomplished by incorporating UV absorbers into the lens materials. It is intuitively obvious that UV radiation up to at least 380nm should be absorbed to provide protection for the cornea, lens and retina. It is also important that the virgin polymer lenses inherently absorb UV more effeciently than clear ophthalmic crown glass lenses (Pitts[53]). UV absorbers incorporated into ophthalmic materials need not discolor the lens. In other words, an almost perfectly clear ophthalmic device can absorb all harmful UVR.

UV absorbing contact lenses have been used in empirical studies to determine if these materials would provide protection against UVR exposure. In recent experiments, pigmented rabbits were fitted with a soft UV absorbing contact lens on one eye and a clear non-absorbing soft contact lens on the other eye. UVR exposure demonstrated very clearly that the UV absorbing lens afforded full protection while the normal soft contact lens showed similar or slightly more damage than an eye not wearing a contact lens (Pitts and Lattimore[54]; Bergmanson et al[55]). This enhanced effect was possibly due to a reduced availability of oxygen through the contact lens but clearly demonstrates that UV absorbing soft contact lens materials do provide protection against UVR exposure.

Summary & Conclusions

The original question was, "Should ocular exposure to solar ultraviolet radiation be a concern?" In an attempt to provide a background to answer this question, several important concepts were introduced. A description of the solar spectral irradiance and the effect of increased and decreased levels of ozone on solar radiation reaching the earth was provided. A short review of the effects of UV exposure to the eye was presented. In addition, the present technology available for protection was reviewed. The answer to the question is quite obvious—yes, solar UVR on earth is of concern, it damages the eye and should be protected against. Table 2 provides a list of persons who should be provided adequate UV protection.

TABLE 2. A list of people who should be protected from exposure to UVR. Item 10 is suggested because of the knowledge that UV exposure is cumulative and in hopes that protection will prevent difficulties for young people and delay those for older adults.

1. Aphakics and pseudo-aphakics to prevent retinal damage.
2. Cataract patients to reduce lenticular scatter.
3. Patients on photosensitizing medication: Chlorothiazides, antibiotics and oral contraceptives are a minimal list of examples.
4. Workers in vocations rich in UVR: Welders, electronic worker, graphic arts workers and researchers.
5. Persons who spend excessive hours in the sun.
6. Patients with pinguecula, pterygia and macular degeneration,
7. Persons enjoying avocations rich in UVR: Snow skiing, sunbathing and mountain climbining.
8. Persons using sunlamps or solariums.
9. Children who play outside or are exposed to excessive UVR to delay photochemical responses in the corneal endothelium, lens and retina.
10. All patients to maintain normal healthy eyes and to eliminate, reduce or delay the prevalence of UV induced corneal problems, cortical senile cataracts and solar retinopathies.

Acknowledgements:

References

1. Widmark EJ, Uber der Einfluss des Lichtes auf die Vorderen Medien des Auge, Skand. Arch Physiol. Vol 1:264-280, 1889.
2. Widmark, EJ, Uber die Durchdringlichkeit der Augenmedien fur Ultraviolette Strahlen, Skand. Arch. Physiol. Vol 3:14-46, 1892.
3. Widmark, EJ, Uber Blendurg der Netzhaut, Skand Arch. Physiol, Vol 4: 281-294, 1893.
4. Verhoeff, FH & L Bell The pathological effects of radiant energy on the eye: an experimental investigation with a systematic review of the literature. Proc Am Acad Arts Sci 51:630-811, 1916.
5. Duke-Elder WS & PM Duke-Elder A histological study of the action of short-waved light on the eye with a note on the 'inclusion bodies'. Br J Ophthalmol 13: 1-37, 1929.
6. Buschke W, JS Friedenwald & SG Moses: Effects of ultraviolet radiation on corneal epithelium: mitosis, nuclear fragmentation, post-traumatic cell movements, loss of tissue cohesion. J Cell Comp Physiol 26:147-164, 1945.
7. Friedenwald JS, W Buschke, J Crowell & A Hollender Effects of ultraviolet irradiation on the corneal epithelium II exposure to monochromatic radiation. J Cell Comp Physiol 32; 161-173, 1948.
8. Cogan, DG and VE Kinsey, Ophthalmic ultraviolet action spectra, Ameri. J. Ophthal 41:969-975, 1956.
9. Rosenthal, FS, C Phoon, AE Bakalian, and HR Taylor, The ocular dose of ultraviolet radiation to outdoor worker, Invest Ophthal and Vis Sci, 29:649-656, 1988.
10. Rosenthal, FS, M Satrum and HR Taylor, The ocular dose of ultraviolet radiation from sunlight exposure, Photochem Photobiol, 42:163-171, 1985.
11. Scotto, J, TR Fears and GB Gori, Ultraviolet exposure patterns, Env. Res. 12:228-237, 1976.

12. Frohlich, C, and RW Brusa, Solar radiation and its variation in time, Solar Physics 74:209-215, 1981.

13. Neckle, H and D Labs, Improved data of solar spectral irradiance from 0.33 to 1.25, Solar Physics, 74:231-249, 1981.

14. Mecherikunnel, AT, JA Gatlin and JC Richmond, Data on total and spectral solar irradiance, Appl Optics, 22:1354-1359, 1983.

15. Green, AES, T Sawada and EP Shettle, The middle ultraviolet reaching the ground, Photochem Photobiol 19:251-259, 1974.

16. Shettle, EP and AES Green, Multiple scattering calculation of the middle ultraviolet reaching the ground, Appl. Optics, 13:1567-1581, 1974.

17. Anon, Stratospheric Ozone Depletion by Halocarbons: Chemistry and Transport, National Academy of Sciences, Chairman of the Panel H. I. Schiff, 1979.

18. Molina, MJ and FS Rowland, Stratospheric sink for chlorofluromethames: chlorine atom-catalyzed destruction of ozone, Nature 249:810-812, 1974.

19. Halpern, P., JV Dave and N Braslau, Sea-level solar radiation in the biological active spectrum, Science 186:1204-1207, 1974.

20. Alyea, FN, DM Cunnold and RG Prinn, Stratospheric ozone destruction by aircraft-induced nitrogen oxides, Science, 188:177-121, 1975.

21. Gille, JC, CM Smythe and DF Heath, Observed ozone response to variations in solar ultraviolet radiation, Science, Vol 22: 315-317, 1984.

22. Kerr, RA, Research News—Taking shots at ozone hole theories, Science 234; 817-818, 1986.

23. Kerr, RA, Research News: Stratospheric ozone is decreasing, science, 239:1489-1491, 1988.

24. Kurtin, WE and JA Zuclich, Action spectrum for oxygen—dependent near-ultraviolet induced corneal damage, Photochem and Photobiol, Vol 27: 329-333, 1978.

25. Ham WT, WA Mueller, JJ Ruffolo, Jr., D. Guerry III, RK Guerry, Action spectrum for retinal injury from near ultraviolet radiation in the aphakic monkey, A.J. Ophthalmol Vol 93:299-306, 1982.

26. Pitts, DG, JE Prince, WI Butcher, KR Kay, RW Bowman, DG Richey, LH Mori, JE Strong and TJ Tredici, The effects of ultraviolet radiation on the eye, SAM-TR-69-10, USAF School of Aerospace Medicine, Brooks AFB, Texas, 1969.

27. Kurtin, WE and JA Zuclich, Action spectrum for oxygen—dependent near-untraviolet induced corneal damage, Photochem & Photobrol, Vol 27:239-333, 1978.

28. Cullen, A.P., Additive effects of ultraviolet radiation, A.J. Optom & Physiol Optics, Vol 57:808-814, 1980.

29. Zuclich, JA Cumulative effects of near-ultraviolet induced corneal damage, Health Phys, Vol 38:833-838, 1980.

30. Zuclich, JA and JS Connolly, Ocular damage induced by near-ultraviolet laser radiation, Invest. Ophthalmol, Vol 15:760-764, 1976.

31. Pitts, DG, JPG Bergmanson, L W-F Chu, M. Waxler and VM Hitchins, Ultrastructural analysis of corneal exposure to UV radiation, Acta Ophthal, Vol 65:263-273, 1987.

32. Cullen AP, BR Chou, MG Hall and SE Jany, Ultraviolet-B Damages Corneal Endothelium, Am. J. Optom & Physiol Optics, Vol 61: 473-478, 1984.

33. Cullen, AP, BR Chou, RF Glover, Corneal endothelial recovery following irradiation with UV-B, Supplement to Invest. Ophthalmol & Vis Sci, Vol 25: 331, 1984b.

34. Doughty, MJ and AP Cullen, Three month evaluation of rabbit corneal physiology following a single exposure to low dose UV-B, ARVO Abstracts, supplement to Invest. Ophthalmol and Vis Sci, Vol 28: 161, 1987.

35. Lattimore, MR, Jr, The effect of ultraviolet radiation on metabolism of the corneal epithelium of the rabbit, PhD Dissertation, College of Optometry University of Houston, August 1987.

36. Blumthaler, M, W Ambach and F Daxecker, On the threshold radiant exposure for keratitis solaris, Invest Ophthalmol & Vis. Sci. Vol 28:1713-1716, 1987.

37. Pitts, DG and AP Cullen, Ocular ultraviolet effects from 295nm to 400nm in the rabbit eye, DHEW (NIOSH) Publication No 77-175, 1977.

38. Hiller R, L. Giacometti and K Yuen, Sunlight and cataract: An epidemiological investigation, A.J. Epi., Vol. 105, 450-459, 1977.

39. Kahn, HA, HM Leibowitz, JP Ganley, MM Kini, T Cohn, RS Nickerson and TR Dawber, The Framingham Eye Study: 1. Outline and Major Prevalence Findings, Am. J. Epidem. Vol 106, 17-32, 1977.

40. Taylor, HR The environment and lens, Brit. J. Ophthal, Vol 64:303-310, 1980.

41. Holows F and D Moran, Cataract—the ultraviolet risk factor. The Lancet, Dec. 5 1249-1251, 1981.

42. Hiller R, RD Sperduto, F Ederer, Epidemiologic associations with cataract in the 1971-1972 national health and nutrition examinations survey 118:239-249, 1983.

43. Leske, MC and RD Sperduto, The epidemiology of senile cataracts: A review, A.J. Epidemiol Vol 118:152-165, 1983.

44. Brilliant LB, NC Grassett, RP Pokhrel, A Kolstad, JM Lepkowski, GE Brilliant, WN Hawks and B Pararajasegaram, Associations among cataract prevalence, sunlight hours and altitude in the Himalayas, A.J. Epidem., Vol 118:250-264, 1983.

45. Pitts, D.G., LL Cameron, JG Jose, S. Lerman, E. Moss, SD Varma, S. Zigler, S. Zigman and J. Zuclich, Optical radiation and cataracts, in Optical Radiation and Visual Health, M. Waxler and V.M. Hitchins, eds, CRC Press, Boca Raton, Fla., pgs. 6-36, 1986.

46. Kamel, ID and JA Parker, Protection from ultraviolet exposure in aphakic erythopsia, Can. J. Ophthalmol Vol 8: 563-565, 1973.

47. Saraux, H, LP Manuet and L LaRoche, Erythopsie chez un porteur d'implant, Etude physiologique et electrophysiologique, J Fr Ophthalmol Vol 7:557-562, 1984.

48. Goodeve, CF, Vision in the ultraviolet, Nature (London) Vol 134:416-417, 1934.

49. Schmidt RE and JA Zuclich, Retinal lesions due to ultraviolet laser exposure, Invest Ophthalmol & Vis Sci, Vol 19:1166-1175, 1980.

50. Henton, WW and SM Sykes, Changes in absolute threshold with light-induced retinal damage, Physiol Behav Vol 31, 179-185, 1983.

51. Henton, WW and SM Sykes, Recovery of absolute threshold with UVA-induced retinal damage, Vol 32:949-954, 1984.

52. Rapp, LM, JG Jose and DG Pitts, DNA repair synthesis in the rat retina following in vivo exposures to 300nm radiation, Invest. Ophthalmol & Vis Sci Vol 26:384-388, 1985.
53. Pitts, DG, Threat of ultraviolet radiation to the eye—how to protect against it, J.A. Optom. Assn., Vol 52:949-957, 1981.
54. Pitts, DG and MR Lattimore, Jr, Protection against UVR using the Vistakon UV-BLOC soft contact lens, ICLC, Vol 14:22-30, 1987.
55. Bergmanson, JPG, DG Pitts and L W-F Chu, The efficacy of a UV-blocking soft contact lens in protecting cornea against UV radiation, Acta Ophthalmol Vol 65:279-286, 1987.

RETINAL IMAGE-MEDIATED OCULAR GROWTH AS A POSSIBLE ETIOLOGICAL FACTOR IN JUVENILE-ONSET MYOPIA
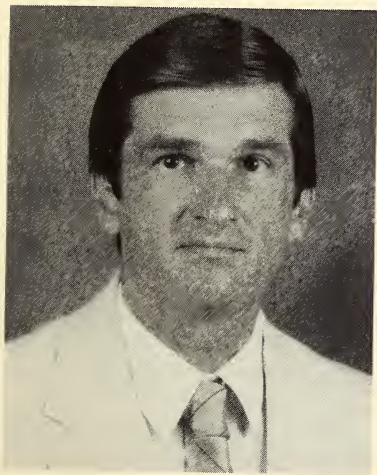
David A. Goss
Professor of Optometry at Northeastern State University College of Optometry, Tahlequah, Oklahoma. From 1984 to 1987 he served as a member of the Working Group on Prevalence and Progression of Myopia, National Academy of Sciences/National Research Council Committee on Vision.

ABSTRACT. Degradation of retinal imagery has been shown to induce myopia in animals by axial elongation through ocular growth. The correlation of the mechanisms of this animal model of myopia and of juvenile onset myopia in humans is not obvious. This paper outlines a hypothesis that bridges this gap and presents a variety of clinical observations consistent with this hypothesis of ocular growth as the mechanism of childhood myopia progression.

RETINAL IMAGE-MEDIATED OCULAR GROWTH AS A POSSIBLE ETIOLOGICAL FACTOR IN JUVENILE-ONSET MYOPIA

INTRODUCTION

Despite the high prevalence of myopia (25 to 30% among young adults in the United States)[1-5] and the considerable attention it has received from vision scientists and clinicians, the determination of its etiologies remains elusive. For instance, the chapter on etiology in Curtin's book on myopia[6] contains 813 citations, but offers no definitive statement on the etiology of juvenile onset myopia. Likewise, treatment regimens designed to slow the progression of myopia have shown negative or inconsistent results.[7-11]

CLASSIFICATION OF MYOPIA

There are undoubtedly several types of myopia with varying etiologies. Necessary to their understanding is an appropriate classification system for the different types of myopia. Grosvenor[12] reviewed existing classification schemes and proposed a new one based on age of onset. In this system the categories of myopia are congenital, youth-onset, early adult-onset, and late adult-onset. Youth-onset or juvenile-onset myopia is the most common form of myopia,[5,12,13] with onset anywhere from about six years of age to the teenage years. Once myopia appears, it increases, or progresses, in amount until its progression stops or slows in the middle or late teens.[14-19] It usually develops to about one to four diopters, and is found with normal corrected visual acuity and normal ocular health. Examples of patterns of childhood myopia progression are shown in Figure 1.

MYOPIA INDUCED IN ANIMALS BY ALTERATION OF OCULAR IMAGERY

One potential technique for investigation of myopia etiology is the development of appropriate animal models. Myopia has been induced in a variety of animal species.[6,20-22]

The common denominator in recent experimental manipulations to induce myopia in animals has been the degradation of retinal imagery.
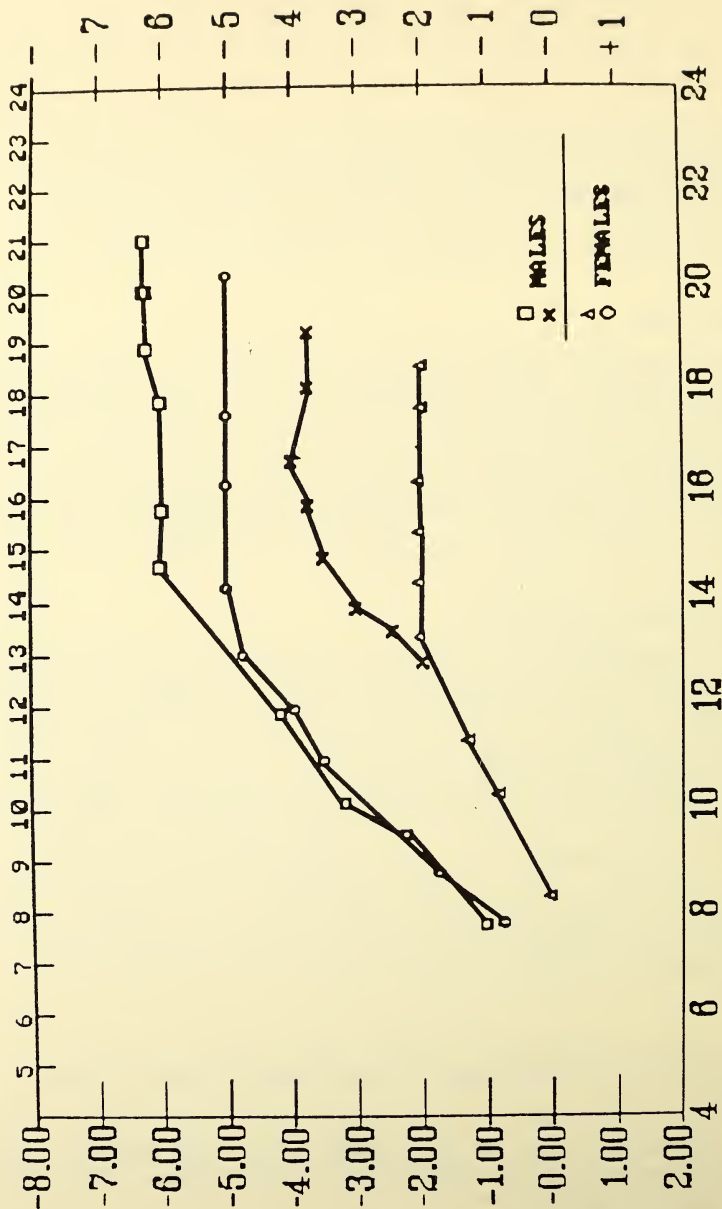
FIGURE 1. Examples of refractive change as a function of age in childhood myopia progression. Refractive error in diopters is on the y-axis, and age in years is on the x-axis. Each set of common symbols represents refractive findings for one patient.

This has been achieved, for instance, by eyelid suture, partial corneal opacification, and attachment of translucent occluders. Myopia has been produced in chickens,[23-28] cats,[29,30] tree shrews,[31] and monkeys.[32-37] In monkeys this line of investigation began with the observation that eyes under sutured eyelids develop axial myopia.[32] Thermal or pressure effects were ruled out when it was found that lid suture did not lead to myopia when the animals were dark-reared,[33] and that corneal opacification induced axial elongation of the eye, and presumably myopia.[34] Smith et al.[35] noted that for monkeys the earlier form deprivation is initiated and the longer it is maintained, the greater is the amount of the myopia that develops.

Animal experimentation suggests that retinal image alteration leads to axial elongation by ocular growth through local mechanisms. Wallman et al.[24] found that when white translucent occluders were applied to selectively deprive the nasal retina or temporal retina of developing chicks, only the deprived part of the eye showed myopia. Vitreous depth increased only on the side of the eye deprived of normal visual input. Optic nerve section does not prevent development of myopia in chicks deprived of form vision.[27]

Local mechanisms are also suggested by the work on rhesus monkeys by Raviola and Wiesel.[37] Various denervation procedures, including ciliary ganglion removal, superior cervical ganglion removal, section of the roots of the trigeminal nerve, optic nerve section, and removal of the striate cortex, did not prevent the development of axial myopia from lid suture. The work of Stone et al.[38] in monkeys may point to a possible mechanism. They found increased levels of vasoactive intestinal polypeptide in retinal amacrine cells in eyes deprived by eyelid suture. Some growth inducing substances, or growth factors, are polypeptides.[39]

The animal experiments described above have produced blur or form deprivation by various forms of translucent occlusion. Schaeffel et al.[40] demonstrated that optical defocus will also lead to myopia in chicks. By application of -4 and -8D lenses they induced a small amount of myopia. The refractive result was approximately the same regardless of whether the lenses were applied monocularly or binocularly. Animals exposed to lenses of +2 and +4 diopters developed a small amount of hyperopia. The posterior nodal distance was greater in eyes treated with minus lenses than in eyes treated with plus lenses. This suggests that axial elongation during development is responsive to the direction of optical defocus. Schaeffel et al.[40] discussed the possibility of accommodation serving as the source of this directional sensitivity.

Human correlates of the high myopia produced in animals by form deprivation exist in humans with eyelid and ocular media anomalies. The severe retinal image degradation caused by lid hemangiomas,[41] ptosis,[42] neonatal lid closure,[43] retinopathy of prematurity,[44,45] and other ocular anomalies[46,47] is associated with a high myopia (generally five diopters or more) in humans. Like the animal model, this form deprivation myopia is due to axial elongation of the eye. (In retinopathy of prematurity there is

both a reduction in anterior corneal radius associated with arrested ocular development and an axial elongation associated with media and retinal changes.) There does not seem to be an obvious correlation of the animal model of myopia with the much more common occurrence of youth onset myopia developing in an otherwise healthy eye. The hypothesis outlined in the next section appears to bridge a gap between form deprivation myopia in animals and clinical observations on juvenile-onset myopia patients.

A NEW HYPOTHESIS FOR THE ETIOLOGY OF JUVENILE-ONSET MYOPIA

Myopia undoubtedly has both genetic and environmental determinants.[48-50] An environmental factor often hypothesized to play a role is near work.[1,3,6,48,49,51-58] Myopic persons often have occupations requiring near work, myopes tend to spend more time reading and in other nearpoint activities, and myopes tend toward better reading ability and scholastic achievement than non-myopes. Myopia is more common in populations that do greater amounts of near work. Though cause and effect have never been definitively demonstrated, a role for near work in myopia development would appear to be a reasonable working assumption.

For clear vision during nearpoint viewing, the dioptric accommodative response must approximate the dioptric stimulus to accommodation. A small lag of accommodation, the dioptric amount by which accommodative response is less than the accommodative stimulus, is normal and will allow clear retinal imagery due to the depth of focus of the eye. Under typical conditions of pupil size, illumination, contrast, and acuity demand, the accommodative response for an accommodative stimulus of 2.50 diopters is approximately 2.25 diopters.[59,60] If the lag of accommodation significantly exceeds this, the quality of retinal imagery may be compromised. Accommodative dysfunction might then lead to myopia by a mechanism similar to that in form deprivation myopia in animals.[50] A lack of contrast in the defocused retinal image might lead to the increased production of growth-inducing substances, or growth factors, which in turn would cause ocular axial elongation.[50,61] In humans, this would most likely be a function of the central retina because foveal vision is used for reading, and objects in the peripheral field are ordinarily not clearly imaged on the retina.[62,63] The slight retinal image degradation associated with accommodative disorders might be expected to produce a lower amount of myopia than that in media opacities; this, of course, is the case in ordinary juvenile- onset myopia. The study by Schaeffel et al.[40] would suggest directional sensitivity of axial growth. That is, if ocular images are in focus at a point behind the retina, growth of the posterior segment would be induced to move the retina in the direction of the focused image. If this is true in humans the optical defocus caused by under-correction of myopia for distance would not result in axial elongation. The direction of optical defocus resulting from insuf-

Age (Yrs)	Hirsch, 1952 (n = 9,552)		Young et al., 1954 (n = 652)		Langer, 1966 (n = 263)	
	Girls	Boys	Girls	Boys	Girls	Boys
5-6	6.15	7.43	4.17	0.00	2.04	0.00
7-8	9.71	11.02	2.60	5.62	3.97	3.08
9-10	17.18	15.68	19.44	9.68	12.20	11.68
11-12	21.60	20.74	20.00	27.27	29.18	20.48
13-14	25.36	22.53	25.71	28.57	34.42	34.30

TABLE 1. Prevalence of myopia in percent as a function of age in school children in the Los Angeles area (Hirsch[64]), in Pullman, Washington (Young et al.[65]), and in Leaside, Ontario, Canada (Langer[66]). The Hirsch and Langer data are the percentages of children with myopia of any amount. Myopia was defined in the Young et al. study as myopia of over one diopter.

ficient accommodation for nearpoint would require axial elongation to achieve focus on the retina.

CLINICAL OBSERVATIONS CONSISTENT WITH THIS HYPOTHESIS

Definitive proof of this hypothesis will be very difficult to obtain. A feasible first step is to examine whether observations on myopia patients are consistent with this hypothesis.

Ages of Onset and Cessation of Childhood Myopia Progression

If the growth factors associated with axial elongation of the eye are synergistic with human growth hormone or somatomedin, then an increase in incidence of myopia would be expected at the times of maximum body growth, and the childhood progression of myopia would be expected to stop or slow when the adolescent growth spurt stops.

Incidence may be inferred from age-related prevalence data. Such data, from three different cross-sectional studies,[64-66] are summarized in Table 1. (Langer's thesis[66] also contained some longitudinal data.) Incidence may be presumed to be greatest at the ages at which prevalence shows the greatest increases. Perusal of Table 1 suggests that the greatest incidence of myopia in girls precedes that in boys in all three populations. The adolescent growth spurt begins about two years earlier in girls than in boys.[67,68]

Once myopia appears, it increases in amount (childhood myopia progression) until the middle or late teens. This can be visualized in Figure 1.

We[15] calculated an index of the age at which childhood myopia progression ceases or slows appreciably. We called this the cessation age, which was calculated using four different graphical and statistical techniques from longitudinal data for several individuals from private practice patient data. One of these methods defined cessation age as the point at which a linear regression of refractive error on age for points below 15 years of age intersected a line of zero slope through the mean amount of myopia at observations after 17 years of age. This was determined individually for patients with three or more refractions before 15 years of age. The mean cessation age was 16.7 years (n = 66, SD = 2.1) for males and 15.2 years (n = 57, SD = 1.7) for females. These are similar to the ages at which the adolescent growth spurt ends.[67,68]

Age of Cessation of Ocular Growth

The progression of myopia in children is due to greater than normal axial elongation of the eye or axial elongation which is inadequately compensated for by reduction in the refractive power of the eye.[19,69] On the basis of the stated hypothesis, it would be expected that ocular growth in myopes would cease at about the same time as the cessation of general body growth and the cessation of childhood myopia progression.

A first glance at the literature would suggest that ocular growth stops before this. Based on cross-sectional and some longitudinal data, Sorsby and his colleagues[70,71] suggested that increases in ocular axial length stop by 13 years of age or before. This inference was based on limited data past 14 years of age, and on a limited number of myopes. Larsen,[72] reporting on his cross-sectional study of axial length, proposed that ocular growth ceases at about 13 years. However, no data after 13 to 14 years of age were reported, and the majority of subjects appeared to be emmetropes and hyperopes.

Thus, childhood myopia progression generally ceases at 15 to 17 years, while literature based largely on non-myopes suggests that ocular growth stops by 13 years of age. A longitudinal study of 18 myopic eyes by Tokoro and Suzuki[73] would seem to indicate that cessation of childhood myopia progression coincides with the cessation of ocular axial elongation. While this question was not specifically addressed in the paper, Tokoro and Suzuki[73] published composite plots of refractive error vs. age and of axial length vs. age. Visual inspection suggests that plateaus occur in both graphs at about the same ages (about 15 to 17 years). To attempt to reconcile this difference, we[74] have collected cross-sectional data on axial length as a function of age from available literature sources and from patients seen in our clinic. We now have a total of 1309 observations. As a preliminary analysis, we calculated the mean axial length as function of age, first for all refractions from all sources, and secondly for myopes from all sources. We calculated linear regression equations of axial length on age using various combinations of consecutive ages with at least ten observations

before the age of 16 years. We determined the regression equation that had the highest coefficient of correlation. The means and regression lines for males with myopia are illustrated in Figure 2. We found cessation ages by determining the points at which the regression equations intersected the mean axial length for observations at and after 17 years of age. The cessation ages thus determined are summarized in Table 2. The cessation ages for increases in stature are also given in Table 2. Mean refractive error, axial length, and height as a function of age are given for male myopes in Figure 2.

When patients of all refractions are considered the derived cessation ages of axial elongation are similar to the ages at which Sorsby and his colleagues and Larsen suggested the eye stops growth. The cessation ages of ocular growth for myopes are later and are close to the mean cessation ages for childhood myopia progression found by Goss and Winkler,[15] with an earlier cessation in females. The cessation ages for ocular growth in myopes are similar to the cessation ages for increase in height.

Effect of Bifocal Lenses on Childhood Myopia Progression

One of the treatment regimens for attempted myopia control is the prescription of bifocal lenses. Reports on the efficacy of this method have shown considerable disagreement.[7,9-11,75,76] One can only say that it is not consistently successful. Potential explanations for the differing results from study to study include investigator bias, the manner in which the bifocals were used, and the types of patients studied.

Work by Roberts and Banford[77] and by Goss[78] suggest that some level of myopia control may be achieved in esophoric patients and patients with accommodative insufficiency. Using data from their patient files, Roberts and Banford calculated rates of progression in diopters per year with age factored out by a formula for the relation of refractive change and age from correlation analysis. The mean rate for single vision lens wearers was -0.405 D/yr, while for bifocal wearers it was -0.314 D/yr. The rates for patients separated by nearpoint phoria are given in Table 3. Rates were less minus for esophores wearing bifocals. They also reported that bifocals reduced the rates for patients with accommodative insufficiency as diagnosed with nearpoint binocular cross cylinder test.

Goss[78] reported on linear regression derived rates of myopia progression (diopters per year) for patient data from three private optometry practices. The mean rate for all single vision lens wearers did not differ significantly from the mean rate for all bifocal lens wearers. For nearpoint esophores the mean rate with bifocals was less than that with single vision lenses (Table 3). Data were also separated by binocular nearpoint cross cylinder net, which was defined as the difference between the gross cross cylinder finding and the binocular maximum plus subjective refraction to best visual acuity. The mean rate with bifocals was less than that with
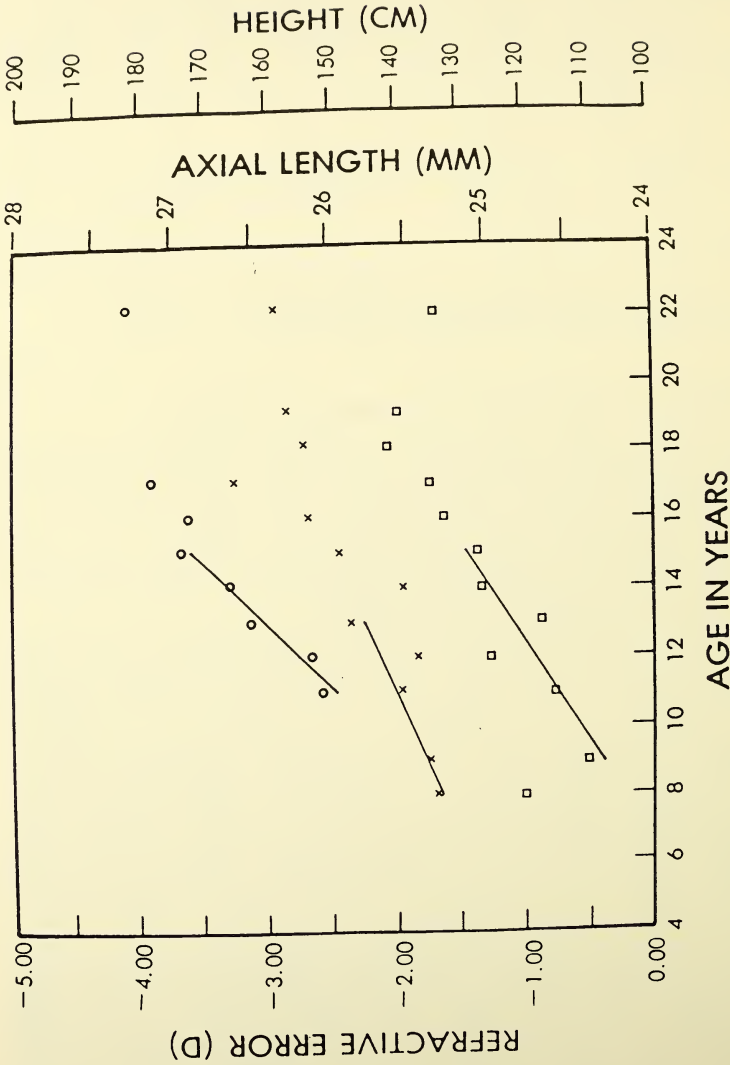
FIGURE 2. Mean axial length in mm (squares), mean refractive error in diopters (X symbols), and mean height in cm (circles) as a function of age in male myopes. Points were plotted for each mean which was determined from at least ten observations. The plotted regression lines are the regressions with the highest coefficient of correlation of each dependent variable with age using various combinations of consecutive age levels up to 16 years. The points plotted at 22 years of age are the means for all observations at 20 years of age and older.

	Males	*Females*
Cessation age of axial elongation in myopes	16.9	14.0
Cessation age of axial elongation in myopes and non-myopes combined	12.7	11.4
Cessation age of increases in height in myopes	16.2	13.8
Cessation age of increases in height in myopes and non-myopes combined	15.7	14.0

TABLE 2. Cessation ages (in years) of ocular growth as assessed by axial length measurements and of general body growth as assessed by height measurements.

single vision lenses for patients with cross cylinder nets of over + 0.50 D (see Table 3).

Morgan[79] gave a normal range for the nearpoint binocular cross cylinder of + 0.25 to + 0.75 diopters over the subjective refraction for distance. Patients with higher plus cross cylinder nets most likely have higher lags of accommodation.[80,81] These are the cases in which the mean rates are lower in magnitude in bifocal wearers. A plus add in bifocal form would improve the quality of retinal imagery in patients with an anomalously high lag of accommodation by reducing the amount of optical defocus. This might then stop the mechanism hypothesized above.

Accommodative disorders prior to onset

Central to the stated hypothesis is that accommodative dysfunction would precede myopia onset. Clinicians have reported observing accommodative disorders in juvenile patients before they became myopic.[82-85] A number of studies[86-91] have reported differences in accommodation and vergence function examined in the laboratory in different types of refractive errors in adults, but this has not been examined quantitatively in children before juvenile onset of myopia. We have begun a study comparing clinical accommodation and convergence findings in children who become myopic and children who remain emmetropic. If this study yields a positive result,

	Single vision lens wearers				Bifocal wearers			Stat. sig. of difference in means
	n	Mean	SD	n	Mean	SD	
Roberts and Banford[77]								
All patients	396	-0.41	----	85	-0.31	----	----
Exo and ortho	181	-0.41	----	17	-0.38	----	----
Esophoria	167	-0.48	----	65	-0.28	----	----
Goss[78]								
All patients	52	-0.44	0.26	60	-0.37	0.24	n.s.
$\geq$ 6 exophoria	9	-0.47	0.31	3	-0.48	0.22	n.s.
0-6 exophoria	27	-0.43	0.21	18	-0.45	0.27	n.s.
esophoria	10	-0.54	0.30	35	-0.32	0.20	p<0.05
cr. cyl. net<0.50	23	-0.31	0.25	6	-0.41	0.22	n.s.
cr. cyl. net$\geq$0.50	16	-0.48	0.30	36	-0.25	0.17	p<0.01

TABLE 3.   Mean rates of childhood myopia progression in diopters per year as a function of nearpoint phoria and nearpoint binocular cross cylinder net (numerical difference between the nearpoint binocular cross cylinder result and the distance binocular subjective refreaction). Standard deviations and statistical significance levels were not given in the original Roberts and Banford publication. Phoria and cross cylinder data were not available for some patients.

additional questions to be asked would be how much defocus is necessary to activate this mechanism and whether appropriate treatment regimens could prevent juvenile-onset myopia.

CONSISTENCY OF THIS HYPOTHESIS WITH A FEEDBACK HYPOTHESIS FOR EMMETROPIZATION

Various hypotheses have been put forward to explain the leptokurtic distribution of refractive errors, with predominance of emmetropia.[92,93] A number of studies have reported inverse correlations of axial length with corneal power and with lens power.[93-97] However, computer simulation by Berck[98] failed to find leptokurtosis of refractive errors as great as observed levels in the population when refractive errors were calculated with correlation of components as observed by Sorsby et al.[95] That is, correlations of components are not sufficient to account completely for emmetropization. Stenstrom[99] and Araki[96] reported leptokurtic distributions of axial length. Sorsby[95] suggested that axial length was normally distributed, but a recent re-analysis of Sorsby's data by Carroll[100] yielded a leptokurtic distribution. Leptokurtosis of ocular axial length may imply that some feed-

back mechanism guides ocular growth. A number of authors [35,37,46,47,101,102] have proposed that the quality of retinal imagery provides such feedback.

SUMMARY AND COMMENTS

Myopia has been induced in laboratory animals by form deprivation. This animal model so far has not been successfully related to the common juvenile-onset myopia in humans. This paper has presented a hypothesis that juvenile onset myopia is related to slightly blurred (defocused) retinal imagery secondary to accommodative dysfunction, through a mechanism of retinal image-mediated ocular growth. While definitive support for this hypothesis is lacking, clinical observations consistent with this hypothesis have been presented. These observations do not rule out other hypotheses, but the hypothesis given here appears to deserve further investigation.

ACKNOWLEDGMENTS

REFERENCES

1. Baldwin, WR. 1964 Some relationships between ocular, anthropometric, and refractive variables in myopia. *Ph.D. Thesis.* Indiana University.

2. Angle, J., and Wissmann, DA. 1980 The epidemiology of myopia. *Am J Epidemiol* 11:220-228.

3. Baldwin, WR. 1981 A review of statistical studies of relations between myopia and ethnic, behavioral, and physiological characteristics. *Am J Optom Physiol Opt* 58:516-527.

4. Sperduto, R. D., Seigel, D., Roberts, J., and Rowland, M. 1983 Prevalence of myopia in the United States. *Arch Ophthalmol* 101:405-407.

5. National Academy of Sciences 1988 *Myopia: Prevalence and Progression.* Committee on Vision Working Group 58. Washington, D.C.: National Academy of Sciences, in press.

6. Curtin, B. J. 1985 *The Myopias: Basic Science and Clinical Management.* Philadelphia: Harper & Row, 61-151.

7. Grosvenor, T. P. 1980 Can myopia be controlled? Parts 1 to 5. *Optom Monthly* 71:504-7, 545-9, 588-91, 620-2, 657-9.

8. Grosvenor, T. P. 1987 Can myopia be controlled? Parts 6 and 7. *Optom Monthly* 72:23-24, 34-36.

9. Goss, D. A. 1982 Attempts to reduce the rate of increase of myopia in young people - A critical literature review. *Am J Optom Physiol Opt* 58:828-841.

10. Young, F. A., Leary, G. A., Grosvenor, T., Maslovitz, B., Perrigin, D. M, Perrigin, J., and Quintero, S. 1985 Houston myopia control study: A randomized clinical trial. Part 1. Background and design of the study. *Am J Optom Physiol Opt* 62:605-613.

11. Goss, D. A., Eskridge, J. B. 1987 Myopia. In: Amos, J. F., ed. *Diagnosis and Management in Vision Care*. Boston: Butterworths, 121-171.

12. Grosvenor, T. 1987 A review and a suggested classification system for myopia on the basis of age-related prevalence and age of onset. *Am J Optom Physiol Opt* 64:545-554.

13. Hirsch, M. J. 1963 The refraction of children. In: Hirsch, M. J., Wick, R. E., eds. *Vision of Children*. Philadelphia: Chilton, 145-172.

14. Hofstetter, H. W. 1954 Some interrelationships of age, refraction, and rate of refractive change. *Am J Optom Arch Am Acad Optom* 31:161-169.

15. Goss, D. A., and Winkler, R. L. 1983 Progression of myopia in youth: Age of cessation. *Am J Optom Physiol Opt* 60:651-658.

16. Goss, D. A., and Cox, V. D. 1985 Trends in the change of clinical refractive error in myopes. *J Am Optom Assoc* 56:608-613.

17. Goss, D. A. 1987 Linearity of refractive change with age in childhood myopia progression. *Am J Optom Physiol Opt* 64:775-780.

18. Goss, D. A. 1987 Matters arising: Cessation age of childhood myopia progression. *Ophthalmic and Physiological Optics* 7:195-196.

19. Goss, D. A. 1989 Characteristics of myopia progression. In.: Grosvenor, T., and Flom, M. C., eds. *Researches on Refractive Anomalies: Clinical Applications*. Boston: Butterworths, in press.

20. Goss, D. A., and Criswell, M. H. 1981 Myopia development in experimental animals - A literature review. *Am J Optom Physiol Opt* 58:859-869.

21. Criswell, M. H., and Goss, D. A. 1983 Myopia development in nonhuman primates - A literature review. *Am J Optom Physiol Opt* 60:250-268.

22. Yinon, U. 1984 Myopia induction in animals following alteration of the visual input during development: A review. *Current Eye Res* 3:677-690.

23. Wallman, J., Turkel, J., and Trachtman, J. 1978 Extreme myopia produced by modest change in early visual experience. *Science* 201:1249-1251.

24. Wallman, J., Gottlieb, M. D., Rajaram, V., and Fugate-Wentzek, L. A. 1987 Local retinal regions control local eye growth and myopia. *Science* 237:73-77.

25. Yinon, U., Rose, L., and Shapiro, A. 1980 Myopia in the eye of developing chicks following monocular and binocular lid closure. *Vision Res* 20:137-141.

26. Rose, L., Yinon, U., and Belkin, M. 1974 Myopia induced in cats deprived of distance vision during development. *Vision Res* 14:1029-1032.

27. Troilo, D., Gottlieb, M.D., and Wallman, J. 1987 Visual deprivation causes myopia in chicks with optic nerve section. *Current Eye Res* 6(8):993-999.

28. Pickett-Seltner, R. L., Weerheim, J., Sivak, J. G., and Pasternak, J. 1987 Experimentally induced myopia does not affect post-hatching development of the chick lens. *Vision Res* 27(10):1779-1782.

29. Seltner, R. L., and Sivak, J. G. 1988 Experimentally induced myopia in chicks. *Can J Optom* 50(3):190-193.

30. Gollender, M., Thorn, F., and Erickson, P. 1979 Development of axial ocular dimensions following eyelid suture in the cat. *Vision Res* 19:221-223.

31. McKanna, J. A., and Casagrande, V. A. 1978 Reduced lens development in lid-suture myopia. *Exp Eye Res* 26:715-723.

32. Wiesel, T. N., and Raviola, E. 1977 Myopia and eye enlargement after neonatal lid fusion in monkeys. *Nature* 266:66-68.

33. Raviola, E., and Wiesel, T. N. 1978 Effect of dark-rearing on experimental myopia in monkeys. *Invest Ophthalmol Vis Sci* 17:485-488.

34. Wiesel, T. N., and Raviola, E. 1979 Increase in axial length of the macaque monkey eye after opacification. *Invest Ophthalmol Vis Sci* 18:1232-1236.

35. Smith, E. L. III, Harwerth, R. S., Crawford, M. L. J., and von Noorden, G. K. 1987 Observations on the effects of form deprivation on the refractive status of the monkey. *Invest Ophthalmol Vis Sci* 28:1236-1245.

36. Thorn, F., Doty, R. W., and Gramiak, R. 1981/2 Effect of eyelid suture on development of ocular dimensions in macaques. *Current Eye Res* 1:727-733.

37. Raviola, E., and Wiesel, T. N. 1985 An animal model of myopia. *New England J Med* 312:1609-1615.

38. Stone, R. A., Laties, A. M., Raviola, E., and Wiesel, T. N. 1988 Increase in retinal vasoactive intestinal polypeptide after eyelid fusion in primates. *Proc Nat Acad Sci* 85:257-260.

39. Beebe, D. C. 1985 Ocular growth and differentiation factors. In Guroff, G., ed. *Growth and Maturation Factors*, vol. 3. New York: Wiley 37-76.

40. Schaeffel, F., Glasser, A., and Howland, H. C. 1988 Accommodation, refractive error, and eye growth in chickens. *Vision Res.* 28(5):639-657.

41. Robb, R. M. 1977 Refractive errors associated with hemangiomas of the eyelids and orbit in infancy. *Am J Ophthalmol* 83:52-58.

42. O'Leary, D. J., and Millodot, M. 1979 Eyelid closure causes myopia in humans. *Experientia* 35:1478-1479.

43. Hoyt, C. S., Stone, R. D., Fromer, C., and Billson, F. A. 1981 Monocular axial myopia associated with neonatal eyelid closure in human infants. *Am J Ophthalmol* 91:197-200.

44. Fledelius, H. C. 1981 Myopia of prematurity-changes during adolescence— a longitudinal study including ultrasound oculometry. *Documenta Ophthalmologica Proceedings* 29:217-223.

45. Goss, D. A. 1985 Refractive status and premature birth. *Optom Monthly* 76:109-111.

46. Rabin, J., Van Sluyters, R. C., and Malach, R. 1981 Emmetropization: a vision-dependent phenomenon. *Invest Ophthalmol Vis Sci* 20:561-564.

47. Nathan, J., Kiely, P. M., Crewther, S. G., and Crewther, D. P. 1985 Disease-associated visual image degradation and spherical refractive errors in children. *Am J Optom Physiol Opt* 62:680-688.

48. Bear, J. C., Richler, A., and Burke, G. 1981 Nearwork and familial resemblances in ocular refraction: A population study in Newfoundland. *Clinical Genetics* 19:462-472.

49. Bear, J. C. 1989 The epidemiology and genetics of refractive error: The present position and feasible advances. In: Grosvenor, T., and Flom, M. C., eds. *Researches on Refractive Anomalies: Clinical Applications*. Boston: Butterworths, in press.

50. Goss, D. A., Hampton, M. J., and Wickham, M. G. 1988 Selected review on genetic factors in myopia. *J Am Optom Assoc* 59(11):875-884.

51. Nadell, M. C., Weymouth, F. W., and Hirsch, M. J. 1957 The relationship of frequency of use of the eye in close work to the distribution of refractive error in a selected sample. *Am J Optom Arch Am Acad Optom* 34(10):523-537.

52. Morgan, M. W. 1967 A review of the major theories for the genesis of refractive state. In Hirsch, M. J., ed. *Synopsis of the Refractive State of the Eye*. A Symposium. Minneapolis: Burgess 8-12.

53. Goldschmidt, E. 1968 *On the Etiology of Myopia - An Epidemiological Study*. Copenhagen: Munksgaard.

54. Borish, I. M. 1970 *Clinical Refraction*, 3rd ed. Chicago: Professional Press, 83-114.

55. Young, F. A. 1977 The nature and control of myopia. *J Am Optom Assoc* 48:451-457.

56. Grosvenor, T. 1977 Are visual anomalies related to reading ability? *J Am Optom Assoc* 48:510-516.

57. Angle, J., and Wissman, D. A. 1978 Age, reading, and myopia. *Am J Optom Physiol Opt* 55:302-308.

58. Richler, A., and Bear, J. C. 1980 Refraction, nearwork, and education - A population study in Newfoundland. *Acta Ophthalmologica* 58:468-478.

59. Ciuffreda, K. J., and Kenyon, R. V. 1983 Accommodative vergence and accommodation in normals, amblyopes, and strabismics. In: Schor, C. M., and Ciuffreda, K. J., eds. *Vergence Eye Movements: Basic and Clinical Aspects*. Boston: Butterworths, 101-173.

60. Ward, P. A. 1987 A review of some factors affecting accommodation. *Clinical and Experimental Optom* 70(1):23-32.

61. Wickham, M. G. 1986 Growth as a factor in the etiology of juvenile-onset myopia. In: Goss, D. A., Edmondson, L. L., and Bezan, D. J., eds. *Proceedings of the 1986 Symposium on Theoretical and Clinical Optometry*. Tahlequah, OK: Northeastern State University 117-137.

62. Duke-Elder, S., and Abrams, D. 1970 *Ophthalmic Optics and Refraction*. In: Duke-Elder, S., ed. *System of Ophthalmology*, vol V St. Louis: Mosby, 141-151.

63. Bennett, A. G., and Rabbetts, R. B. 1984 *Clinical Visual Optics*. London: Butterworths 298-303.

64. Hirsch, M. J. 1952 The changes in refraction between the ages of 5 and 14. Theoretical and practical considerations. *Am J Optom Arch Am Acad Optom* 29:445-459.

65. Young, F. A., Beattie, R. J., Newby, F. J., and Swindal, M. T. 1954 The Pullman study - A visual survey of Pullman school children. *Am J Optom Arch Am Acad Optom* 31:111-21, 192-203.

66. Langer, M. A. 1966 Changes in ocular refraction from ages 5-16. *M.S. Thesis.* Indiana University.

67. Tanner, J. M. 1973 Growing up. *Scientific American* 229(3):34-43.

68. Lowrey, G. H. 1978 *Growth and Development of Children.* 7th ed. Chicago: Yearbook Publishing.

69. Sorsby, A. 1980 Biology of the eye as an optical system. In: Safir, A., ed. *Refraction and Clinical Optics.* Hagerstown, MD: Harper & Row 133-149.

70. Sorsby, A., Benjamin, B., and Sheridan, M. 1961 *Refraction and its Components During the Growth of the eye from the Age of Three*, Medical Research Council Special Report Series No. 301, London: Her Majesty's Stationery Office.

71. Sorsby, S., and Leary, G. A. 1970 *A Longitudinal Study of Refraction and its Components during Growth*, Medical Research Council Special Report Series No. 309, London: Her Majesty's Stationery Office.

72. Larsen, J. S. 1971 The sagittal growth of the eye. IV. Ultrasonic measurement of the axial length of the eye from birth to puberty. Acta Ophthalmologica 49:873-886.

73. Tokoro, T., and Suzuki, K. 1969 Changes in ocular refractive components and development of myopia during seven years. *Japanese J Ophthalmol* 13:27-34.

74. Goss, D. A., Cox, V. D., Porter, G., Nielsen, E., and Dolton, W. No Date Refractive error, axial length, and height as a function of age in myopes, manuscript in preparation.

75. Grosvenor, T., Perrigin, D. M., Perrigin, J., and Maslovitz, B. 1987 Houston myopia control study: A randomized clinical trial. Part II. Final report by the patient care team. *Am J Optom Physiol Opt* 64:482-498.

76. Hemminki, E., and Parssinen, O. 1987 Prevention of myopic progress by glasses. Study design and the first year results of a randomized trial among school-children. *Am J Optom Physiol Opt* 64:611-616.

77. Roberts, W. L., and Banford, R. D. 1967 Evaluation of bifocal correction technique in juvenile myopia. *Optom Weekly* 58(38):25-8, 31;58(39):21-30;58(40):23-8;58(41):27-34; 58(43):19-24,26.

78. Goss, D. A. 1986 Effect of bifocal lenses on the rate of childhood myopia progression. *Am J Optom Physiol Opt* 63(2):135-141.

79. Morgan, M. W. 1944 Analysis of clinical data. *Am J Optom Arch Am Acad Optom* 21(12):477-491.

80. Fry, G. A. 1940 Significance of the fused cross cylinder test. *Optom Weekly* 31(1):16-19.

81. Borish, I. M. 1970 *Clinical Refraction.* Chicago: Professional Press.

82. Avetisov, E. S. 1979 Unterlagen zur Entstehungstheorie der Myopie. 1. Mitteilung. Die Rolle der Akkommodation in der Entstenhung der Myopie. *Klinische Monatsblatter Augenheilkunde* 175:735-740.

83. Avetisov, E. S. 1980 Unterlagen zur Entstehungtheorie der Myopie. 4. Mit-
     teilung. Entstehung der Myopie und einige neue Moglichkeiten zu ihrer
     Prophylaxe und Therapie. *Klinische Monatsblatter Augenheilkunde* 176:911-914.

84. Birnbaum, M. H. 1979 Management of the low myopia pediatric patient. *J
     Am Optom Assoc* 50:1281-1289.

85. Birnbaum, M. H. 1981 Clinical management of myopia. *Am J Optom Physiol
     Opt* 58:554-559.

86. Maddock, R. J., Millodot, M., Leat, S., and Johnson, D. A. 1981 Accom-
     modative responses and refractive error. *Invest Ophthalmol Vis Sci* 20:387-391.

87. McBrien, N. A., and Millodot, M. 1986 Amplitude of accommodation and
     refractive error. *Invest Ophthalmol Vis Sci* 27:1187-1190.

88. McBrien, N. A., and Millodot, M. 1987 The relationship between tonic ac-
     commodation and refractive error. *Invest Ophthalmol Vis Sci* 28:997-1004.

89. Bullimore, M. A., and Gilmartin, B. 1987 Aspects of tonic accommodation
     in emmetropia and late-onset myopia. *Am J Optom Physiol Opt* 64:499-503.

90. Rosenfield, M., and Gilmartin, B. 1987 Synkinesis of accommodation and
     vergence in late-onset myopia. *Am J Optom Physiol Opt* 64:929-937.

91. Bullimore, M. A., Boyd, T., Mather, H. E., and Gilmartin, G. 1988 Near
     retinoscopy and refractive error. *Clinical and Experimental Optom* 71(4):114-
     118.

92. Everson, R. W. 1973 Age variation in refractive error distributions. *Optom
     Weekly* 64(9):31-34.

93. Goss, D. A. 1988 Theories of emmetropization: a review. In: Goss, D. A.,
     and Edmondson, L. L., eds. *Proceedings of the 1988 Northeastern State University
     Symposium on Theoretical and Clinical Optometry*. Tahlequah, OK: Northeastern
     State University, 90-98.

94. Stenstrom, S. 1948 Investigation of the variation and the correlation of the
     optical elements of human eyes - Part IV, Woolf, D., translator. *Am J Optom
     Arch Am Acad Optom* 25(8):338-397.

95. Sorsby, A., Benjamin, B., Davey, J. B., Sheridan, M., and Tanner, J. M. 1957
     Emmetropia and its Aberrations. Medical Research Council Special Report Series
     No. 293. London: Her Majesty's Stationery Office.

96. Araki, M. 1962 Studies on refractive components of human eye by means of
     ultrasonic echogram. Report III. The correlation of among refractive com-
     ponents. *Acta Societatis Ophthalmologica Japonicae* 66(2):128-147.

97. van Alphen, G. W. H. M. 1961 On emmetropia and ametropia. *Ophthalmo-
     logica* 142 (supplementum): 1-92.

98. Berck, R. R. 1983 An examination of refractive error through computer sim-
     ulation. *Am J Optom Physiol Opt* 60(1):67-73.

99. Stenstrom, S. 1948 Investigation of the variation and the correlation of the
     optical elements of human eyes - Part III, Woolf, D., translator. *Am J Optom
     Arch Am Acad Optom* 25(7):340-350.

100. Carroll, J. P. 1980 Geometrical optics and the statistical analysis of refractive error. *Am J Optom Physiol Opt* 57(6):367-371.

101. Wallman, J., Adams, J. I., and Trachtman, J. N. 1981 The eyes of young chickens grow toward emmetropia. *Invest Ophthalmol Vis Sci* 20(4):557-761

102. Medina, A. 1987 A model for emmetropization: predicting the progression of ametropia. *Ophthalmologica* 194:133-139.

VISUAL EVOKED POTENTIAL (VEP) MEASURES OF VISION IN UNILATERAL APHAKIC INFANTS

Daphne L. McCulloch
Director of the Electrophysiological Unit in the Department of Ophthalmology at the Hospital for Sick Children, Toronto, Ontario, and Assistant Professor of Ophthalmology at University of Toronto.

ABSTRACT. A longitudinal study of 16 infants with unilateral deprivational amblyopia was conducted using VEPs. The infants fell into 3 clinical groups: 1) those with dense cataracts diagnosed early in life, 2) those with dense cataracts along with other ocular complications, and 3) those with partial cataracts. All infants had early surgery, contact lens correction and an aggressive patching regimen to treat their amblyopia.

All infants in Group 1 developed useful vision in their amblyopic eyes. VEPs showed rapid improvement in the first 8 to 10 months post correction. However, a residual deficit in the aphakic eye remained throughout the study. Group 2 infants, with complicated histories, had a dichotomous pattern of visual recovery. Some showed excellent results comparable to those in Group 1. Others showed little or no visual recovery. Infants with

partial cataracts, Group 3, had an initial milder amblyopia and progressed to better visual outcomes than all other infants.

This study has demonstrated that early surgical treatment, contact lens correction and patching leads to an excellent prognosis for significant visual recovery from unilateral deprivational amblyopia.

VISUAL EVOKED POTENTIAL (VEP) MEASURES OF VISION IN UNILATERAL APHAKIC INFANTS

Complete unilateral deprivation of patterned vision early in life induces the most severe form of amblyopia[1,2]. In human infants, deprivational amblyopia can result from conditions such as dense congenital cataract, persistent hyperplastic primary vitreous (PHPV) and ptosis. Deprivational amblyopia presents a more serious clinical challenge than the relatively milder forms of strabismic and anisometropic amblyopia. In addition, deprivational amblyopia forms a much closer analogy to animal models of amblyopia induced by lid suture or occlusion.

The prognosis for development of visual resolution in animals monocularly deprived from birth improves if the deprivation period is short and if the fellow eye is occluded soon after deprivation is ended. Studies of cats suggest that optimal improvements in acuity without inducing amblyopia in the fellow eye are achieved by patching the fellow eye 50% to 70% of the time[3].

Following unilateral visual deprivation from dense congenital cataracts, deficiencies in a number of visual functions have been reported in unilateral aphakic infants and children. These include: decreased visual resolution, asymmetric optokinetic nystagmus, loss of peripheral sensitivity (particularly in the nasal field), loss of contrast sensitivity (particularly to high spatial frequency stimuli), diminished VEP responses and poor colour discrimination [4, 5, 6, 7, 8, 9, 10, 11, 12, 13]. There is general agreement that, in humans as well as in animals, early correction and patching of the fellow eye markedly improves visual function. Asymmetric optokinetic nystagmus is an exception which does not respond to patching[10].

Most studies of infants born with unilateral visual deprivation have used psychophysical techniques to measure visual functions[5, 6, 7, 10, 12, 13, 14]. Only a few cases of congenital monocular visual deprivation have been studied using VEP assessments of visual function[8, 9]. The medical histories and management varies considerably among these reported cases.

A group of 16 infants with deprivational amblyopia have been followed longitudinally using VEPs. All infants had visual input restored after early surgical removal of the cataract and/or vitreous opacity and aphakic contact lens correction. Similarities in medical histories and management of this

group allow a general description of VEP development following early visual deprivation.

Subjects

Since 1984, the aphakic infants who had surgical and contact lens management at The Hospital for Sick Children have been referred to the visual electrodiagnostic unit for serial VEP assessments. Of the 54 infants and children who have been assessed, 16 met the following two criteria for inclusion in this study: 1) They had monocular visual deprivation resulting in a unilateral lensectomy prior to 16 months of age, 2) They have been followed in the unit for at least 1 year and have had at least 4 assessments. Table 1 lists all of the subjects and gives the principle characteristics of each case.

The 16 children were divided into 3 diagnostic groups. Group 1 (7 infants), had uncomplicated unilateral congenital cataracts which completely blocked patterned vision and allowed only diffuse light to reach the retina. To be included in group 1, infants must have had a presurgical description of their cataract made prior to 6 months of age which included a drawing or description of a dense central cataract, a notation of dense leucocoria and/or the notation that the ocular fundus was completely obscured. Infants in group 1 had normal visual systems except for their unilateral cataract and any of the three conditions most commonly associated with cataract: 1) strabismus (3 infants had esotropia, two of these were corrected surgically during the assessment period); 2) microphthalmus and microcornea (Post surgical ultrasound measurements of axial length differed in 5 out of 7 cases with a range of 1.5 to 2.5 mm of difference. Corneal diameters differed by 0.5 mm or more in 6 of 7 cases.); and 3) latent nystagmus was present in all aphakic eyes during early assessments. (Latent nystagmus improved during the assessment period in 3 cases and resolved in 4 cases).

All infants in group 1 received contact lens corrections prior to 8 months of age. To be included in group 1 infants must have patched their phakic eye at least 25% of the waking hours.

Six infants also had complete unilateral visual deprivation as described above but they had significantly different histories which could affect their prognosis. These 6 infants were considered separately as group 2 and included 3 cases of dense persistent hyperplastic primary vitreous (PHPV) one of these cases had an associated macular retinal fold and diagnosis after 5 months of age, 1 case of congenital rubella syndrome with hearing loss, mental retardation and "salt and pepper" retinal pigmentation, 1 case with latent nystagmus in both the phakic and aphakic eyes, and 1 case in which there was no compliance with the recommended patching regimen.

A third group, 3 infants, were monocularly deprived of clear vision but may have had the ability to resolve at least coarse shapes during the deprivation period. All of group 3 infants were diagnosed after 5 months

TABLE 1: Description of Subjects

Subject	History	Age at diagnosis (weeks)	Age at lensectomy (membranectomy)	Age at CL dispensing	Age at strabismus surgery
Group 1					
KA	cataract OS	4.0	7.6	10.4	-
JB	cataract OS	5.0	5.5	10.7	-
JE	cataract OS, hyaloid remainant	0.3	1.5	4.9	-
LK	cataract OS	3.0	16.4	19.2	-
DP	cataract OD, secondary membr.	4.3	4.4 (26.2)	10.1 (37.9)	48
AS	cataract OS	10.8	14.2	18.3	-
AT	cataract OS	21.0	24.1	34.6	88
Group 2					
EM	cataract OD no patching	5.4	6.6	18.8	-
DG	PHPV OS	8.8	25.6	30.4	-
BM	PHV OD	4.3	10.8	13.9	-
AP	PHPV OD retinal fold	34.6	38.1	54.6	79
AH	cataract OD, rubella syndrome	17.6	45.3	50.8	-
AM	cataract OD, bilateral nystagmus	13.0	13.9	19.4	-
Group 3					
BA	post. lenticonus OD	22.0	33.6	36.8	111
KH	post. lenticonus OD, & partial cataract	39.0	48.1	55.2	147
GR	partial cataract OD	33.3	67.5	76.6	-

of age. Two infants had posterior lenticonus, one of these had a partial cataract as well. The third infant had a posterior polar cataract which did not block the fundus view. None of the infants in group 3 had latent nystagmus although microphthalmus (2 of 3) and esotropia (2 of 3) were present in this group.

Clinical Management

In order to minimize deprivational amblyopia, the clinical management had three basic objectives: 1) Early surgery with removal of the visual obstruction 2) Prompt and accurate refractive correction with contact lenses and 3) Patching of the non-deprived eye for 50% of the waking hours. All infants in this study had a unilateral lensectomy with anterior vitrectomy.

Shortly after surgical recovery, each infant had an examination under anesthesia (EUA) which included fundus examination, measurement of axial length, keratometry, and a trial contact lens fitting of the aphakic eye. Both soft and rigid gas permeable contact lenses were used with the lens providing the best positioning and over refraction selected on an individual case basis. Visual deprivation was considered to have ended when the first contact lens was dispensed for daily wear. Contact lenses were fitted to allow clearest vision at a 50 centimeter distance (i.e. over refraction of -2.00 diopters).

Group 1, the infants with uncomplicated complete cataracts, had the earliest diagnosis, surgery and contact lens dispensing (median ages of 4.3, 7.6 and 10.7 weeks respectively). Infants with partial visual deprivation, group 3, were treated later than those in all other groups principally because they were diagnosed later (5 to 9 months of age at diagnosis). The ages of diagnosis, surgery and contact lens dispensing for each individual case are given in table 1.

The contact lens histories of these aphakic infants include many brief interruptions for lost lenses or illness as well as many refittings to correct for the refractive and fitting changes associated with rapid growth of the eye early in life. All of these patients were monitored very regularly in the contact lens clinic, as often as necessary to establish a satisfactory correction and every three months thereafter. Brief periods of interruption or inappropriate refractive correction were not considered as periods of visual deprivation for the purpose of this study. However, one infant who required secondary membrane surgery was considered visually deprived until after a clear visual axis was restored.

Compliance with the 50% patching regimen was good across all three groups of patients with a few exceptions. One infant in each group (DP, AH and GR) had moderate compiance (13 to 37%). Subject EM with right aphakia did not patch and was placed in group 2 for this reason. Subject AP with PHPV and a macular retinal fold was not advised to patch her unaffected eye because of the poor prognosis for visual recovery in the aphakic eye but she continued patching a few hours per week.

Regular VEP assessments were introduced as part of the routine clinical care of these aphakic infants. Feedback was given to the parents and reports were sent to the medical and contact lens files. However, results of VEP assessments were not used to modify any other aspect of clinical management of these subjects.

Methods

VEP assessments began with a brief vision screening in which pupillary responses, fixation and contact lens over refraction were checked.

Standard VEP recordings to pattern reversing checkerboard stimuli were measured using a Nicolet CA2000 clinical evoked potential system. The active electrode was placed one centimeter above the inion with the

reference and ground electrodes on the earlobes. Electrode impedance was below 5 mega ohms. The infant was seated 75 cm in front of a video screen which generated checkerboard patterns of 95% contrast switching at 1.88 reversals per second. Incoming voltages were fed to preamplifiers then filtered with a bandpass of 0.1 to 35 Hz. The VEPs were averaged over a 400 millisecond epoch beginning at the stimulus reversal. Twenty to 40 epochs were averaged for each trial.

Several modifications facilitated the testing of infants and young children. The infants were seated on their parent's lap in a darkened booth with no visible distractions. An observer located behind the stimulus monitored fixation and operated an interrupt switch so that VEPs were collected only while the infants were watching the stimulus. This observer also made sounds and jingled keys in front of the stimulus to maintain interest. An artifact rejection function was engaged which eliminated all epochs contaminated by head and body movements.

Infants were tested monocularly with an adhesive eye patch used as an occluder. Check size could be varied from 7.5 minutes of arc to full screen flash in single octave steps as follows: 7.5, 15, 30, 60, 120, 240, 480 minutes and flash. At least two trials of each check size were attempted with each eye, except that sizes above 120 minutes were not tested when VEP responses were clearly present to smaller patterns, and small patterns were not tested when there was no response to medium sizes. Cooperation was rated excellent, good, fair or poor by the observer, based on the infant's attentiveness or fussiness.

An attempt was made to have the infant wear a compensating lens in a trial frame whenever the contact lens overrefraction indicated more than 3 diopters of blur at the stimulus distance (i.e. an over refraction greater than $+1.75$ or less than -4.50). VEP assessments were scheduled approximately monthly while the infants were under one year of age, every 3 months up to 2 years of age and every 6 months thereafter.

Data Analysis

The morphology of the transient VEP changes rapidly in the first 6 months of life[15, 16] and is quite different in normal and amblyopic eyes[17]. For these reasons no attempt was made to define specific VEP peaks or troughs. VEPs were defined as present or absent to a specific stimulus based on repeatability. That is, the VEP is present if one or more "substantial" peaks or troughs occurred at the same latency (± 20 msec.) on both trials to that stimulus size. To avoid measuring small baseline fluctuations, a "substantial" peak or trough was defined as the largest peak to trough amplitude within a 130 milliseconds segment (± 65 msec.) of that trial. Analyses were based on a) the minimum check size which produced a repeatable VEP and b) the latency and amplitude of the VEP to the 120 minute stimulus whenever this response was present. Latency was meas-

ured to the first repeatable peak and the greatest peak to trough amplitude within a 130 millisecond time was used as a measure of amplitude.

Results

A total of 124 assessments were completed on 16 monocularly aphakic children (median number of assessments per child is 8 with a range of 4 to 15). Assessments were conducted when the children were aged 10 weeks through 3 years, 9 months. Only two assessments were aborted due to uncooperative infants. VEP traces from a typical assessment are illustrated in figure 1. Examples of the longitudinal changes in minimum check size eliciting a VEP and the in latency of the VEP to the 120 minute stimulus are shown in figure 2.

In group 1, all 7 infants showed similar patterns of VEP responses. Specifically, the minimum check size which elicited a VEP in the aphakic eye decreased sharply over the first 8 to 10 months (34 to 47 weeks) post deprivation with a small residual deficit in the aphakic eye continuing up to 2 1/2 years (152 weeks) post correction. Figure 3 shows the combined data for group 1. Small improvements in the phakic eyes are apparent from the presurgical point through the first 6 weeks post correction. This is the improvement associated with normal rapid development in VEP responses in the first 6 months of life[18, 19]. Development of the phakic eyes cannot be followed in this study because most infants have a response to the smallest stimulus (7.5') by 5 months of age.

The VEP to a large patterned stimulus (120') is markedly delayed or absent in the earliest post correction assessments of children in group 1. Thereafter, the aphakic eyes show a small delay in VEP latency and much larger intersubject variability than the latency of the VEPs from phakic eyes. Figure 4 shows the average latencies for all 7 cases in group 1.

VEP amplitudes showed extreme variability both within and among subjects. To decrease variability, VEP amplitudes were expressed as the relative amplitudes of the VEP from the aphakic eye compared with the phakic eye within the same session. VEPs from aphakic eyes had generally smaller amplitudes than those from normal eyes (see figure 5). However, even this measure of relative amplitude is too variable to allow monitoring of individual cases.

Since group 2 infants have more varied histories, their data were not combined. Two general patterns were found in the VEP assessments of the group 2 infants. Three infants (DG, AH, and AM) showed post-deprivation VEP development which was indistinguishable from the infants in group 1. (16 of 17 assessments of the aphakic eyes of these three infants fell within 1 standard deviation of the averages for group 1 as shown in figures 3 and 4.) Data for the minimum check size to elicit VEPs from two of these cases are shown on the right of Figure 6. AM differs from other cases because she had latent nystagmus in both eyes. Her phakic eye showed poorer responses than the normal eyes of other infants.

Right   Left

120'

60'

30'

15'

Monocular Aphakia
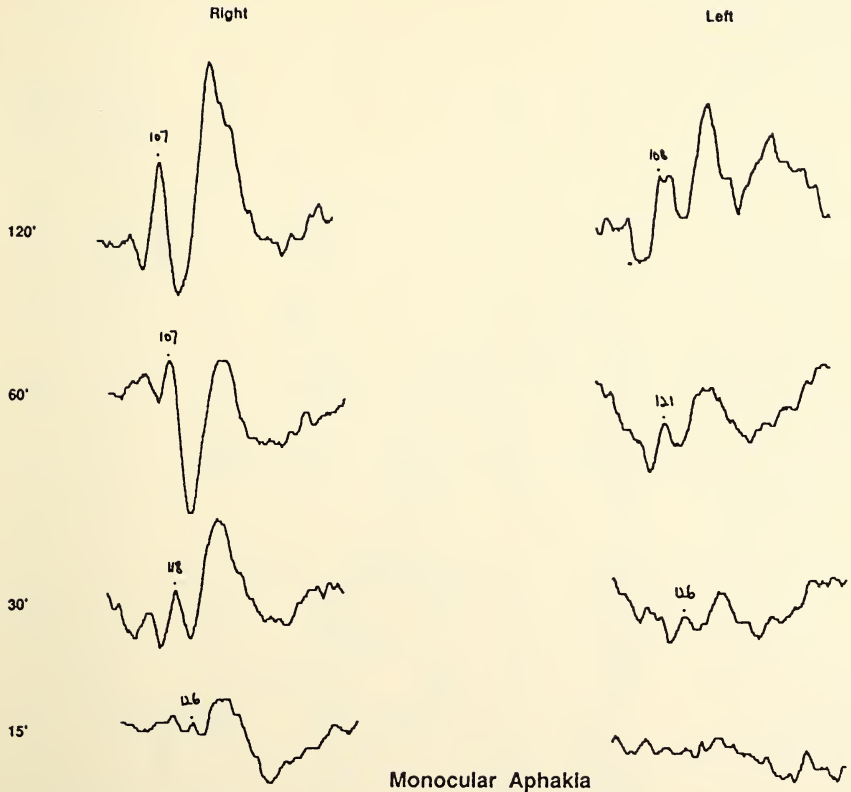
Figure 1: VEP traces from one assessment of subject LK are shown above for the right phakic eye and the left aphakic eye. Stimuli were checkerboards alternating at 1.88 reversals per second. Check sizes in one octave steps from 120 to 15 minutes of arc were used. The latency to the first positive peak is labeled for each peak which showed repeatability of individual trials. The assessment shown above was conducted when LK was 50 weeks (11 months) old, 30 weeks after she received a contact lens correction in her left eye.

Daphne L. McCulloch
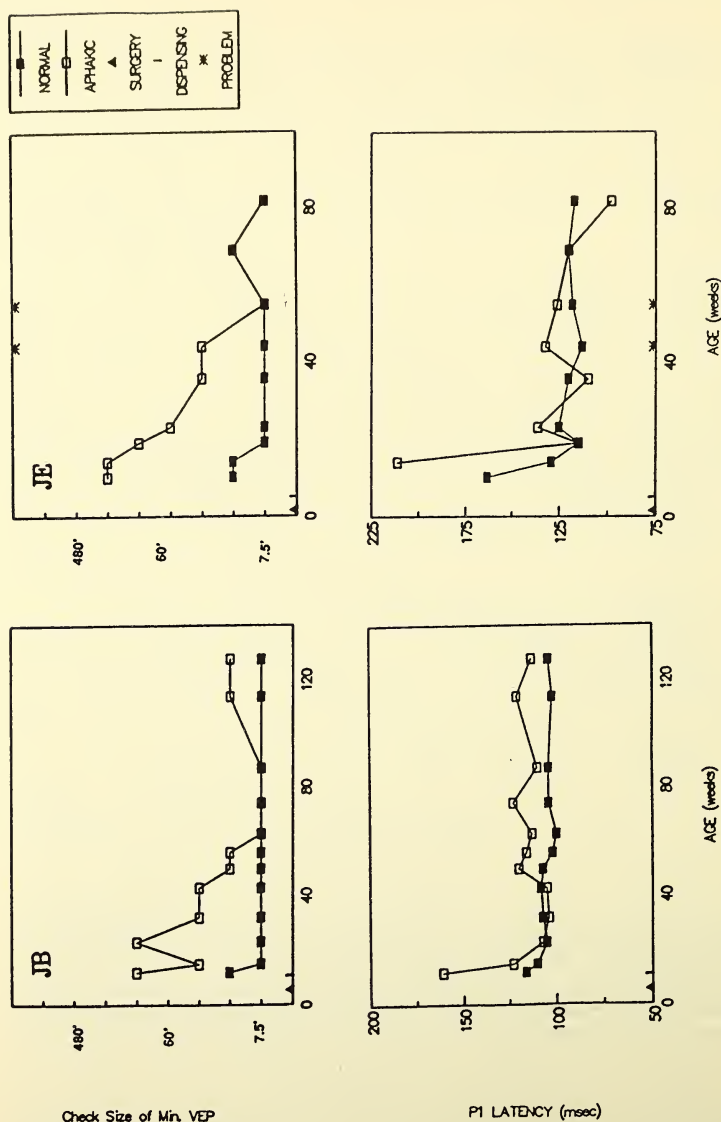
Figure 2: Data from serial assessments of subjects JB and JE are shown. On the upper graphs the smallest check size which elicited a repeatable VEP is plotted versus age at each VEP assessment. The latency of the first rpeatable posative peak of the VEP elicited by the 120' stimulus is shown on the lower graphs. Both of these subjects had dense congenital cataracts in their left eyes with correction early in life (see Table 1). The asterisk (*) indicates assessments with poor cooperation or with more than 3 diopters of refractive overcorrection at the test distance.
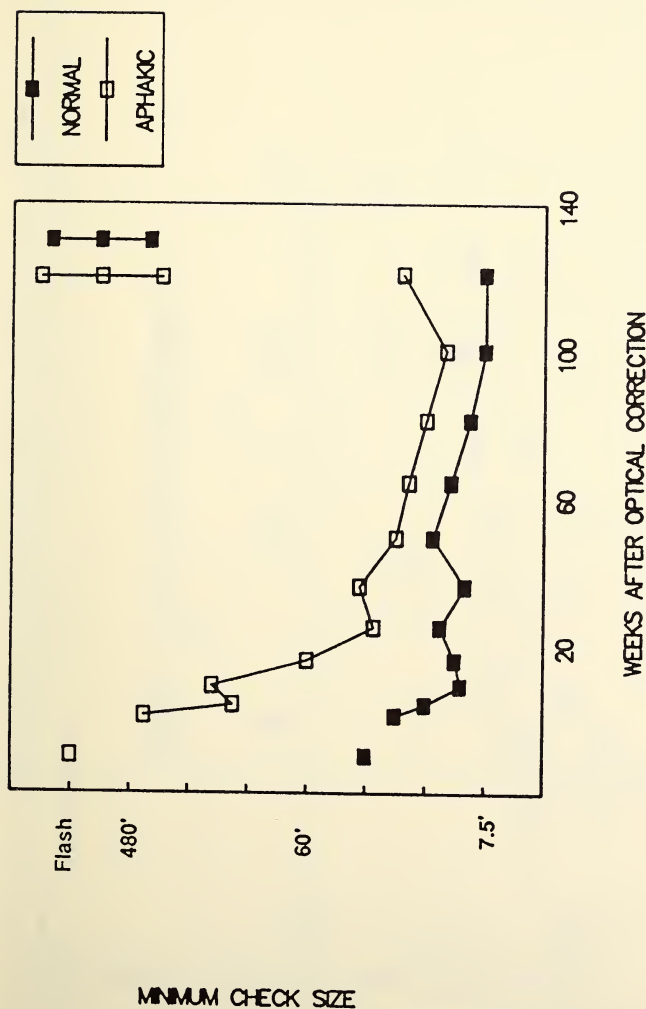
Figure 3: Mean values for the minimum check size eliciting a repeatable VEP are shown for the 7 infants with uncomplicated dense unilateral cataracts (group 1). Means were calculated on a logarithmic scale and the time scale was normalized by placing all assessments into bins according to the number of weeks after the patient received the first contact lens correction. Flash responses were placed on the ordinate at 920′, the width of the stimulus screen. Cumulated standard deviations for phakic and aphakic eyes are shown in the upper right corner of the graph. The first points on the right of the graph are the results of presurgical assessments (n = 2). Other points are means of about 6 assessments (range 4-9 from a total of 63 assessments).

Daphne L. McCulloch

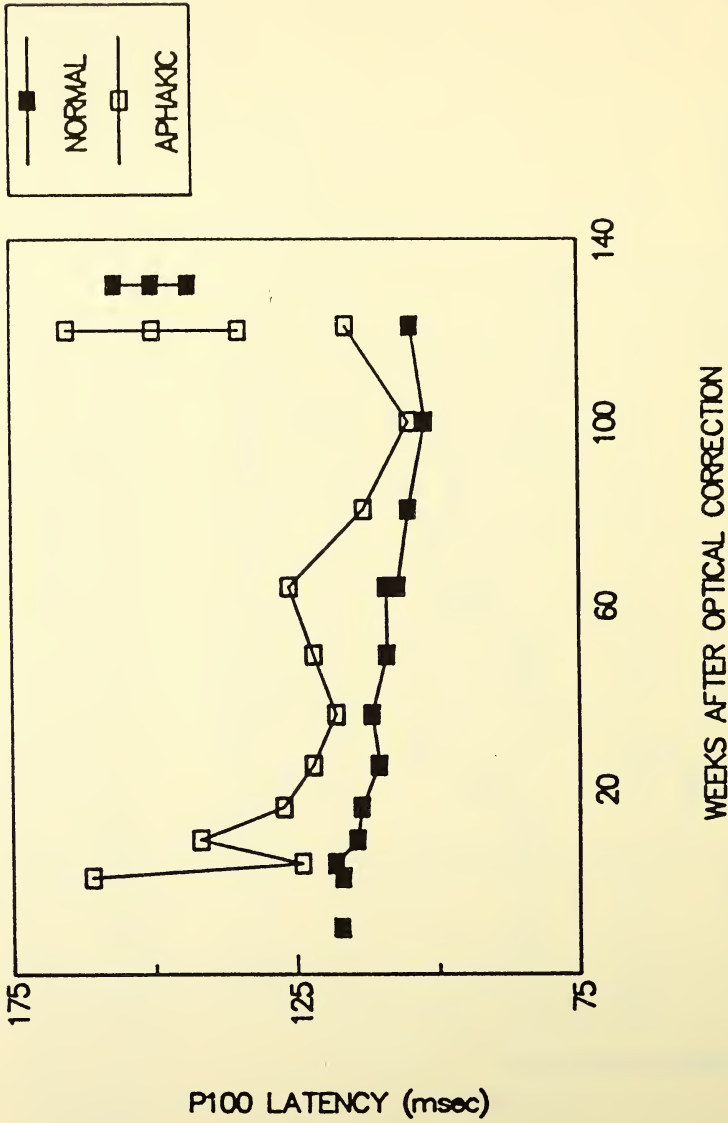**Figure 4:** Mean values for the P100 latency of the VEP elicited by the 120' check stimlus are shown for group 1 infants (n = 7). As in Figure 3, the time scale is normalized by setting the time of contact lens correction to zero. Presurgical assessments and the earliest post correction assessments do not have a latency for the aphakic eye because there was no VEP response to the 120' stimulus.

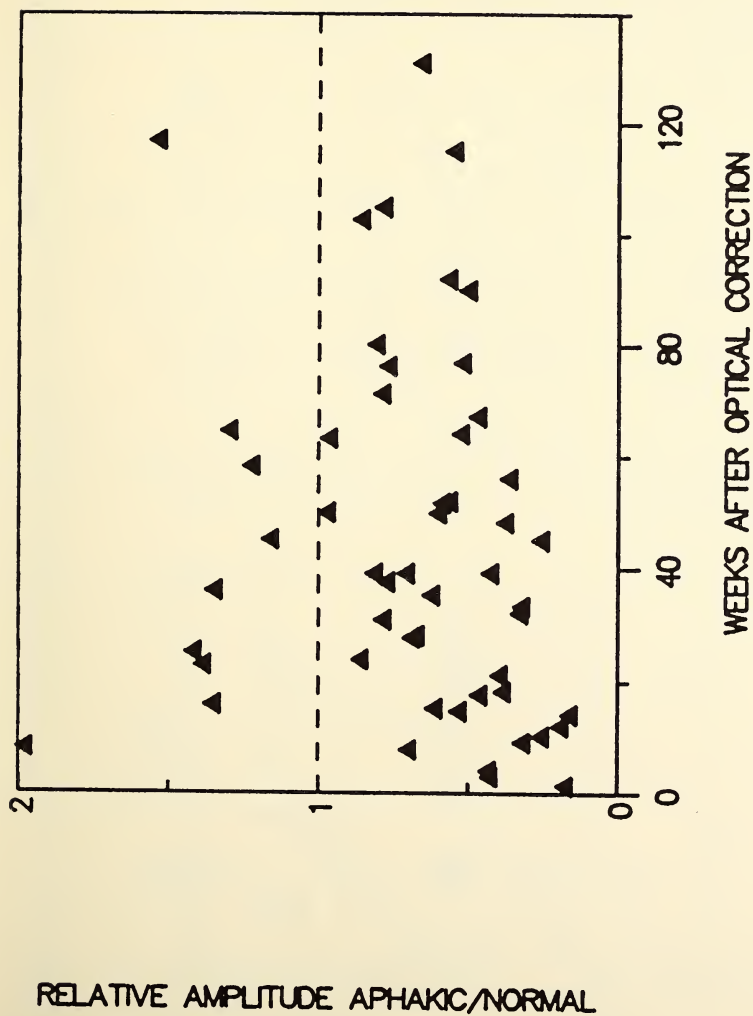**Figure 5:** The VEP from the aphakic eye is shown as a fraction of the amplitude of the VEP from the phakic eye for the VEP elicited by the 120' stimulus. Ten assessments with no response from the aphakic eye to the 120' stimulus are not shown. The majority of points (44/53) are below 1 indicating that the VEP from the aphakic eye usually has a smaller amplitude than the VEP from the phakic eye.
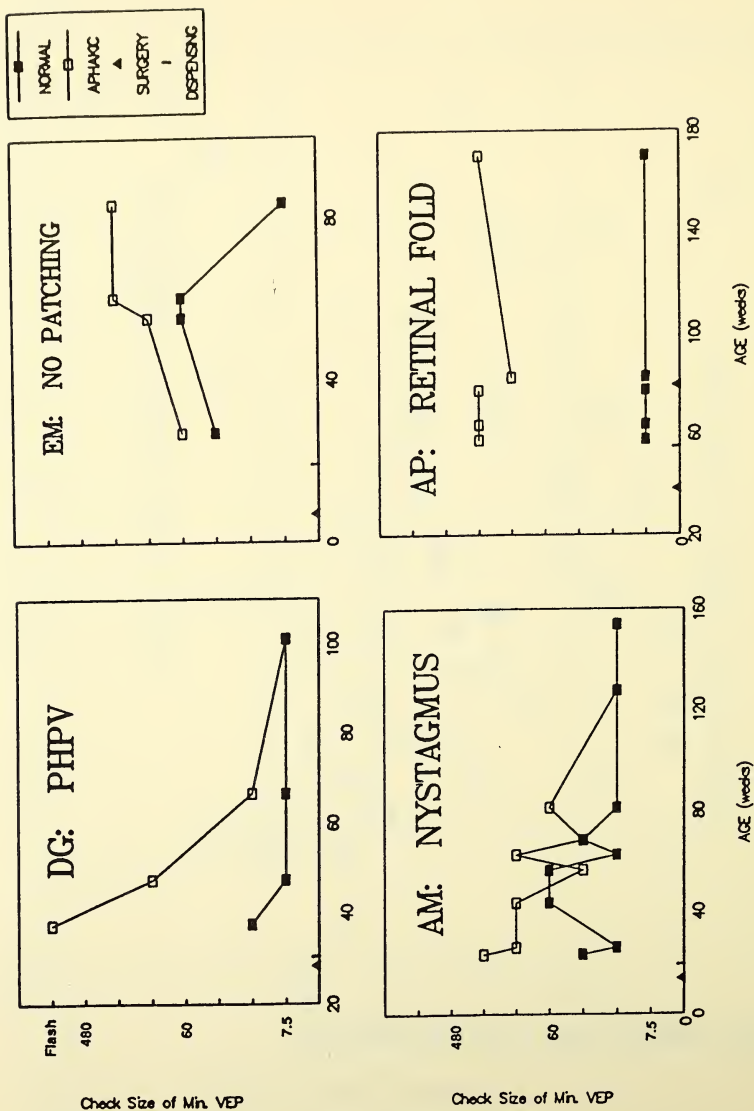
Figure 6: Minimum check size to elicit a VEP is plotted versus age for four subjects in group 2. The aphakic eyes of D.G. and A.P. showed poor or no development of the VEP to small check sizes following optical correction. See text and table 1 for the case histories of these subjects.
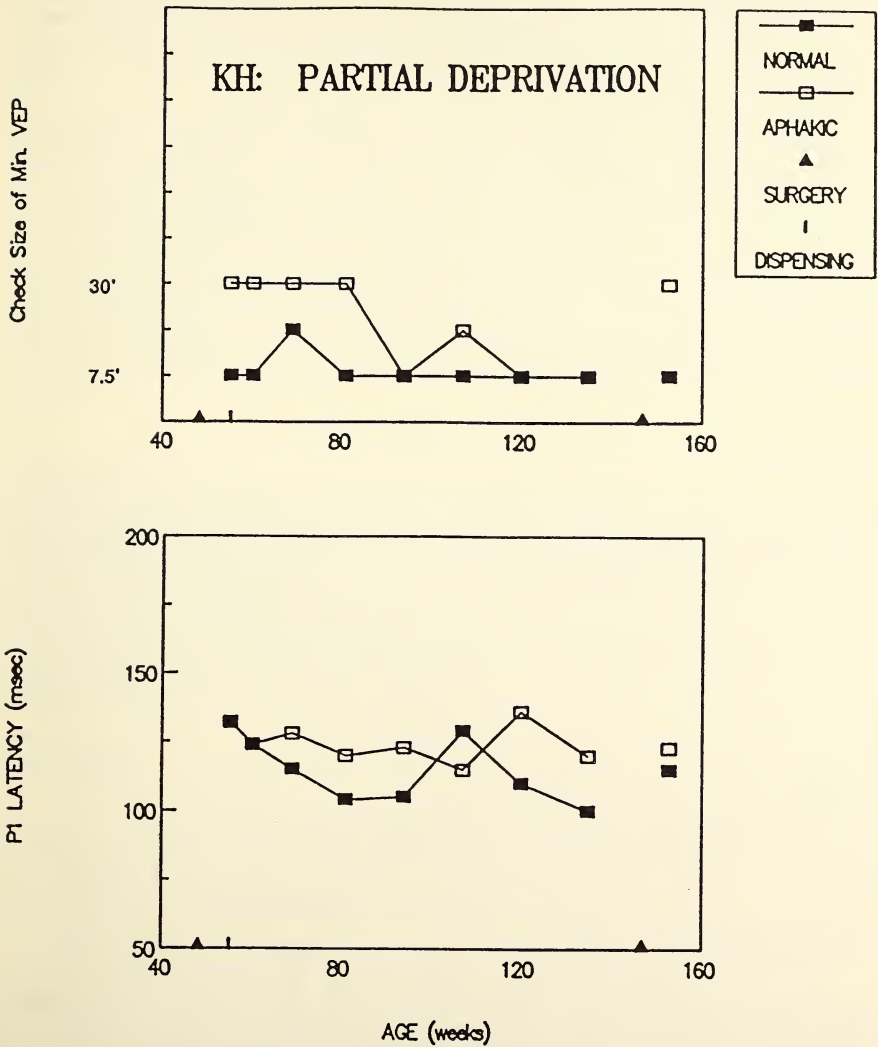
Figure 7. The minimum check size eliciting a VEP and the P100 latency to the 120' check stimulus is shown for subject KH who was born with a partial cataract in her right eye. The first surgery indicated was her lensectomy, the second was a strabismus repair.

The other three cases in group 2 (EM, BM and AP) showed little or no improvement in their aphakic eyes following deprivation (see the left graphs in figure 6). One of these cases (EM) is particularly notable because her only visual "complication" was that she did not comply with the patching regimen. She did, however, wear a contact lens correction. The other two cases with poor development were AP who had a macular retinal fold and BM who had PHPV.

All three infants with partial visual deprivation, group 3, had better VEPs from their aphakic eyes than the infants with complete deprivation, even though these infants received later diagnoses and correction. Figure 7 shows the minimum check size which elicited a VEP in KH who was typical of this group. All 3 of these children are now old enough to have visual acuities measured by the Sheridan-Gardiner method. They have mild to moderate amblyopia (acuity range 6/12 to 6/36) but a VEP response was often elicited by our smallest stimulus, 7.5 minutes of arc checks, in their aphakic eyes thus reaching the sensitivity limit of this analysis.

Conclusion

All seven infants with uncomplicated dense unilateral cataract showed rapid visual development in the first 8 to 10 months after correction. Exceptional children who developed reasonably good vision following complete visual deprivation have been reported in previous studies[4, 10, 13]. However, this series has shown that useful visual function can be expected when visual deprivation is treated with early surgery, early correction and occlusion of the undeprived eye for 50% of the waking hours. Our case, EM, who did not occlude showed almost no visual development in her aphakic eye. This supports the finding of other authors,[10, 12, 13] that occlusion can be crucial for post-deprivation visual development even when there has been early surgery and correction.

Several cases illustrate good post-deprivation visual development when PHPV, congenital rubella syndrome and bilateral nystagmus were present along with unilateral visual deprivation. However, the visual prognosis in PHPV may be poorer than in cataract. Two cases of PHPV, one with retinal complications, did not show significant post-deprivational development.

The age, at correction of the infants with complete unilateral deprivation ranged from 5 to 37 weeks. Although the group is not large enough for rigorous statistical evaluation, it is interesting to note that within this young group there were no obvious differences in either the rate of visual development or the eventual VEP threshold of the aphakic eye. Children treated for dense congenital cataracts at later ages can have a markedly poorer visual outcome (e.g. correction after 6 months[10], correction at 11 and 21 months [5]). However, these older patients did not receive consistent patching so the effect of prolonged deprivation alone cannot be distinguished from the effect of shorter deprivation without patching (see also[13]).

Since 5 to 37 weeks of visual deprivation produces amblyopia which is partially reversible, it is clear that the critical period for visual development extends at least up to 37 weeks of age. The deprivational amblyopia is not completely reversible so it would seem that the critical period for normal visual development has begun by 5 weeks of age. However, some caution must be introduced in interpreting the residual visual deficits as entirely due to irreversible deprivational amblyopia. Nearly all infants in this study had microphthalmus in their aphakic eye. These small eyes cannot be assumed to have the same visual potential as fully developed eyes. In addition, aphakic correction does not restore entirely normal visual input. Aphakic eyes experience blur at some distances because they cannot accommodate and partial deprivation from strabismus or from uncorrected refractive errors are frequently present beyond the initial period of complete visual deprivation. Studies of aphakic children with deprivational amblyopia are continuing. As numbers are increased it is hoped that the understanding of visual deprivation and reversal of its effects will be increased further.

Aknowledgements

References

1. Boothe R.G. 1981 Development of spatial vision in infant macaque monkeys under conditions of normal and abnormal visual experience. Pp. 217 - 242 in Aslin, Alberts and Peterson, eds., Development of Perception. Volume. 2. New York: Academic Press.
2. Mitchell D.E. & Timney B. 1983 Postnatal development of function in the mammalian visual system. Pp. 507 - 917 in J.D.Smith, ed. Handbook of Physiology: the nervous system III, Baltimore. Williams & Wilkins.
3. Mitchell D.E. 1987 Recovery of visual function after early monocular deprivation in kittens: Implications for Treatment of Human Amblyopia. Symposium on visual development, Meeting of visual scientists of Toronto, Toronto, October 1987.
4. Enoch, J.M. & Rabinowicz, I. 1976 Early surgery and visual correction of an infant born with unilateral lens opacity. Doc. Ophthal. 41(2) 371-382.
5. Hess R.F. France T.D. Tulunay-Keesey V. 1981 Residual vision in humans who have been mocularly deprived of pattern stimulation in early life. Exp. Brain Research 44: 295-311.

6.   Jacobson S.G. Mohindra I. Held R. 1981 Development of visual acuity in infants
     with congenital cataracts. British Journal of Ophthalmology 65: 727-735.
7.   Jacobson S.G. Mohindra I Held R. 1983 Monocular visual form deprivation in
     human infants. Doc. Ophthal. 55: 199-211.
8.   Levi D.M. Harwerth R.S. 1980 Contrast sensitivity in amblyopia due to stim-
     ulus deprivation. Br.J. Ophthal. 64: 15-20.
9.   Odom V.J. Hoyt C.S. Marg E. 1981 Effect of natural deprivation and unilateral
     eye patching on visual acuity in infants and children. Arch. Ophthalmol
     99:1412-1416.
10. Lewis T.L. Maurer D. Brent H.P. 1986 Effects on perceptual development of
     visual deprivation during infancy. British Journal of Ophthalmology.70:214-
     220.
11. Mioche L Perenin M.T. 1986 Central and peripheral residual vision in humans
     with bilateral deprivation amblyopia. Experimental Brain Research 62: 259-
     272.
12. Maurer D. Lewis T.L. Brent H.P. 1988 The effects of deprivation on human
     visual development: Studies of children treated for cataracts in Morrison F.J.,
     Lord C.E. and Keating, D.P. Applied Development Psychology Vol. 3.
13. Tytla M.E. Maurer D. Lewis T.L. Brent H.P. 1988 Contrast sensitivity in chil-
     dren treated for congenital cataract. Clinical Vision Sci. 2: 251-264.
14. Atkinson J. & Braddick O. 1982 Assessment of visual acuity in infancy and
     early childhood. Acta. Ophthal. 157 (Suppl),18-26.
15. Sokol S. 1982 Infant visual development: evoked potential estimates. Ann.
     N.Y Acad.Sci. 388: 514-523.
16. Taylor M. Meinzies L. MacMillan L. Whyte H. 1987 VEPs in normal fullterm
     and premature neonates: longitudinal versus cross-sectional data. Electroenc.
     & Clin. Neurophys. 68: 20-27.
17. Arden G. & Barnard W. 1979 Effect of occlusion on the VER in amblyopia.
     Trans of the Ophth. Soc. UK. 99: 419-426.
18. Atkinson J. 1984 Human Visual Development over the first 6 months of age.
     A review and hypothesis. Human Neurobiol. 3: 61-74.
19. Boothe R. Dobson V. Teller D. 1985 Postnatal development of vision in human
     and nonhuman primates. A. Rev. Neurosci. 8: 495-545.

ANALYTICAL GEOMETRY OF THE VISUAL FIELD
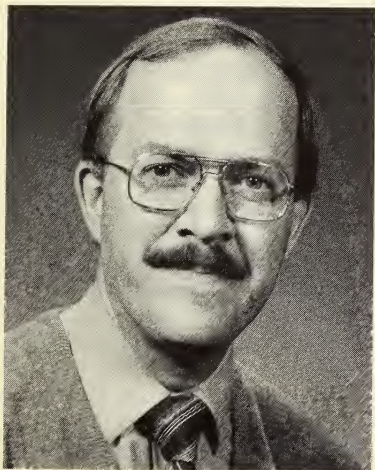
T. David Williams
Professor of Ocular Pathology at the
School of Optometry, University of
Waterloo, Waterloo, Ontario.

ABSTRACT. Many clinical decisions regarding the visual field are based on a qualitative analysis of visual field charts. This paper describes ways to make visual field data (specifically, Goldmann perimeter charts) accessible to the computer. This permits analysis of visual field area, which in turn makes it possible to do a quantitative analysis of effects of age and pupil diameter on the size of the visual field. Other uses of this analysis described in the paper include: an assessment of the size of visual field quadrants as a percentage of the total field area; earlier recognition of abnormal isopters in glaucoma; differing effects of glaucoma and retinitis pigmentosa on the visual field area; and differences among phakic and aphakic populations.

ANALYTIC GEOMETRY OF THE VISUAL FIELD

A great deal of the information generated in an optometric setting is on paper or in a photographic form. Analysis of this information, for example a Goldmann perimetry chart, is frequently qualitative: the clinician makes a decision based on clinical experience. If ways can be found to transfer such information into a more quantitative form, then this analysis may enable the clinician and researcher to make more informed decisions.

Earlier researchers have used a planimeter to trace the outline of isopters and calculate the area they contain.[1] A more current method of analysis may be performed by taking the object (photograph or visual field chart) to be analysed and placing it on a digitizing pad. By lining up a cross-hair or stylus with pertinent points, it is possible to digitize (that is, generate X and Y coordinates for) each point on the chart. These coordinates may be used in conjunction with the formulas of analytic geometry to calculate various attributes of the two-dimensional object under analysis. This method was previously used to measure the effects of provocative testing with topical cortisone in patients with ocular hypertension.[2] I have used this sort of system to calculate the tortuosity of retinal arterioles and the shape of the optic nerve head (on fundus photographs) and to fit ellipses to visual field isopters and blindspots.

Digitizing a Visual Field Chart

If one places a Goldmann visual field chart on the digitizing table and determines X and Y coordinates for the points around a given isopter, as well as for the fixation point, it is possible to calculate the area contained within the isopter, the area contained within a given quadrant of the visual field, or the area within a given sector of the visual field. The visual field areas discussed in this paper are expressed in terms of the area (in square millimeters) contained within a given isopter (or quadrant) on the surface of the Goldmann visual chart. Digitizing the isopter permits the visual field information to be manipulated numerically: this allows a more quantitative analysis.

Effects of Aging on the Visual Field

Clinicians have long been aware that the extent of the visual field of an older patient is usually smaller than that of a younger patient.

Fig. 1 shows the results of digitizing 130 Goldmann I-2 isopters for a group of people 10 to 75 years of age.[3,4] Only one isopter for each person

Figure 1. Scatter plot of the I-2 visual field area for the right eye of 130 normal patients aged from 10 to 75 years. The three dashed lines represent (from top to bottom) the upper 95% confidence limit for the mean area at any age, the least squares line, and the lower 95% confidence limit for the mean area at any age. Reprinted by permission from Williams TD. Aging and visual field area. Am J Optom Physiol Opt 1983;60(11):888-891.

was used for this analysis. These individuals were all free of ocular or systemic disease. The Goldmann I-2 test target yields a visual field equivalent in size to that obtained with a 1 mm white target at the 1 meter tangent screen.

The least squares line fitted to these data is $y = -83x + 8616$. The correlation for these data $(r) = 0.72$. From the least squares line, one may calculate that the Goldmann I-2 isopter area decreases at a rate of about 10% per decade. A 75-year-old individual would be expected to have a field area equal to 25% of the y-intercept value, or 2154 square millimeters. The 75-year-old in the study had an I-2 field area of 2118 square millimeters.

Fig. 2 shows results of a quadrant-by-quadrant analysis of the above-mentioned visual fields. The number of isopters for each age group is shown in Fig. 3: in this case, right and left visual field results of each individual were used. Again, as clinical experience would suggest, we see that the temporal quadrants are larger than the nasal quadrants, and also that the inferior quadrants tend to be larger than the superior quadrants. Data for the Goldmann I-2 and I-4 isopters are shown in Fig. 2. The I-4 test target is 1 log unit more intense than the I-2 test target. Despite the large changes in visual field area over the 4 decades shown, the quadrants
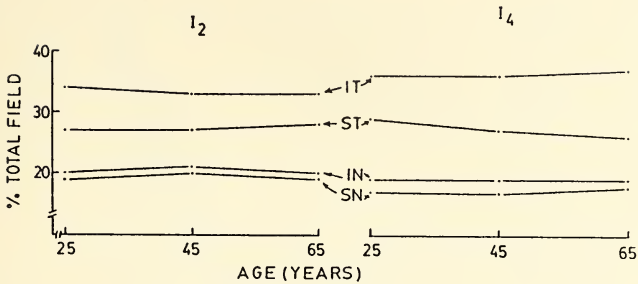
Figure 2. Area of the visual field quadrants for the same group of people as Fig. 1, expressed as a percentage of the total visual field area, as a function of age. Despite a roughly 40% decrease in the total visual field area over the 4 decades shown, each quadrant occupies a remarkably constant proportion of the total visual field. Data shown in Fig. 2 are drawn from right and left eye field data for these people: number of isopters and ages are as shown in Fig. 3.

occupy the same percentage of the total field. This finding suggests that all quadrants of the visual field are equally affected by the aging process.[5]

Effects of Increasing Stimulus Intensity

Fig. 3 shows the effect of intensity on the area of the visual field for the group of individuals described above. For this figure, right and left isopter data were combined. On the X-axis, the stimulus value is obtained by adding together the Roman and Arabic numerals of the Goldmann stimuli; thus, the I-2, I-3, and I-4 stimuli have stimulus values of 3, 4 and 5 respectively.

The slope of the functions for the age groups shown is the same. It appears that, while older individuals have smaller visual fields, their ability to respond to increased stimulus intensity is unchanged. This trait of the visual field may be used to identify patients with various ocular diseases.

Fig. 4 shows data for a 53 year old patient with chronic open-angle glaucoma. Goldmann visual field charts are shown in Fig. 4A and B. Testing was done with the I-2, I-3 and I-4 stimuli. The visual field for the left eye (Fig. 4B) shows a profound loss of sensitivity in the superior nasal quadrant. At first glance, the field for the right eye (Fig. 4A) appears unremarkable. Analysing the isopter areas, one sees in Fig. 4C that the left eye's response to increased stimulus intensity is different from the norm for that age group. The right eye's response in Fig. 4C is at the lower end of the normal range. If the superior nasal quadrant areas are analysed, the results shown in Fig. 4D are obtained: these show a nearly flat response as far as the left eye is concerned, but the response of the right eye is departing more and more
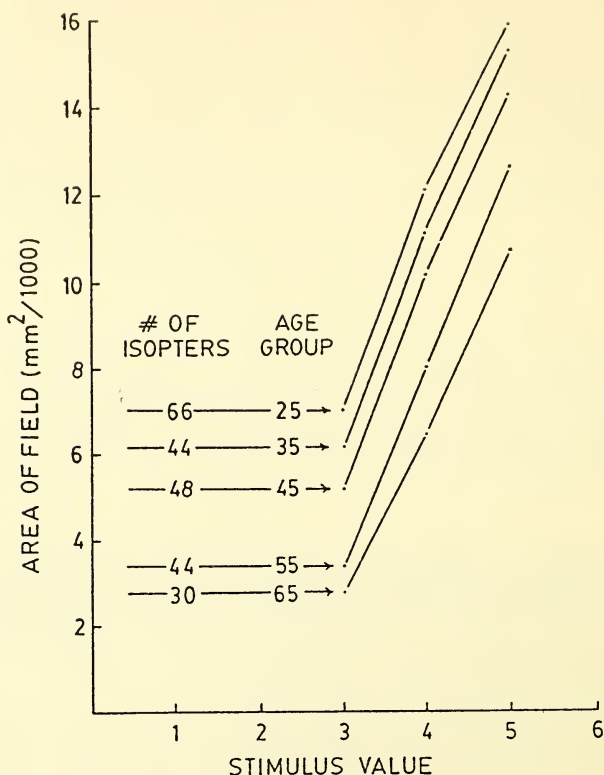
Figure 3.   Effect of increasing stimulus intensity on the area of the Goldmann visual field. Stimulus value is determined by adding together the Roman and Arabic numbers which specify the target area and intensity; thus, the stimulus values for the I-2, I-3, and I-4 targets are 3, 4, and 5 respectively. Reprinted by permission from Williams TD. Aging and the isopter area. Ophthalmic Optician 1985;25(4):98-102,

from the normal as stimulus intensity is increased.[6] The slope of these functions provides a further quantitative value for analysis of visual field data.

Fig. 5 shows the response of a patient with retinitis pigmentosa to increased stimulus intensity. In this case, it appears that the departure from normal decreases if the stimulus intensity is increased sufficiently (note the equal slopes for the last limb of the RP function and the normal function). While the data of Fig. 4 were not obtained with the same wide range of stimulus intensities as those in Fig. 5, I would suggest that the
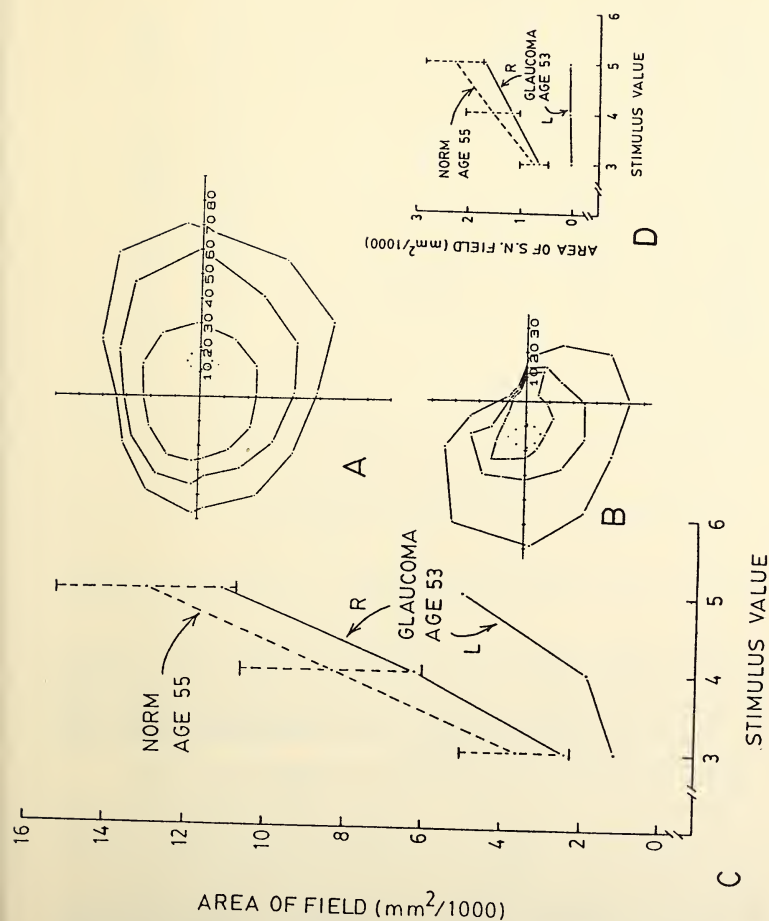
Figure 4. Goldmann visual field charts for right eye (A) and left eye (B) of a 53 year old with chronic open-angle glaucoma. C: area analysis for both sets of isopters shows abnormal response of left eye to increased stimulus intensity, and low-normal response of right eye to increased stimulus intensity. D: area analysis for superior nasal quadrant of right eye shows that visual field area falls farther below normal as stimulus intensity is increased. Reprinted by permission from Williams TD. Visual Field Area Response to Increased Target Intensity: A Method of Detecting Ocular Disease. Am J Optom Physiol Opt 1986;63(1):28-31.
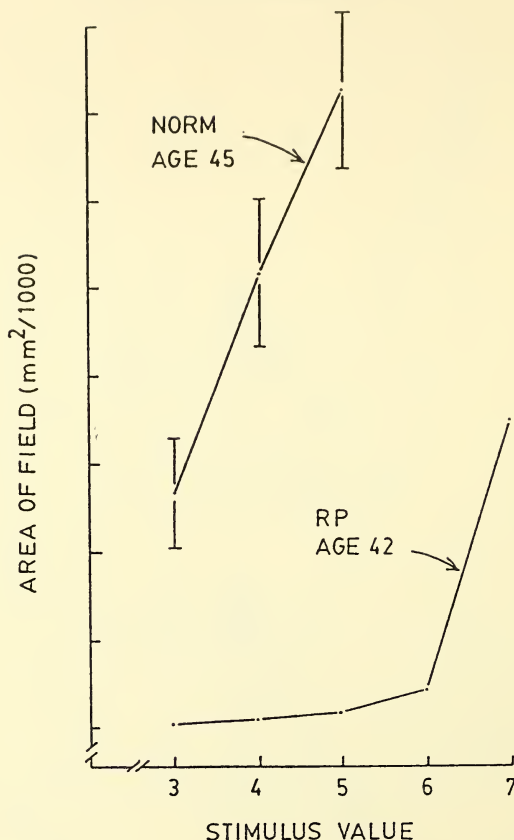
Figure 5. Visual field area data for 27 normal patients aged 40 to 50 years (upper function). Error bars show 1 SD on either side of the mean values. Visual field area for five isopters of a 42-year-old female with retinitis pigmentosa are shown on lower function. Note that the slope of this function more closely approximates normal when stronger stimuli are used. Reprinted by permission from Williams.[6]

visual fields for these two different disease conditions show different responses to increased stimulus intensity. It may be that conditions affecting receptors show a more normal field area vs intensity slope when the intensity is increased, while conditions which affect (for example) ganglion cells show a less normal field area vs intensity slope when intensity is increased. This question clearly requires more investigation.
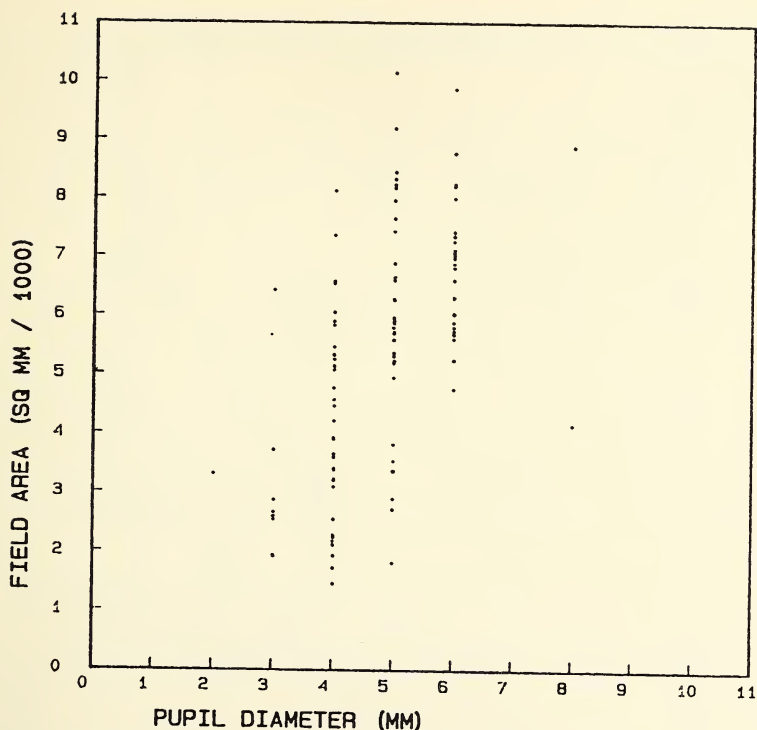
Figure 6. Scatterplot of I-2 visual field area and pupil diameter for 130 normal individuals between 10 and 75 years of age.

Causes for Decreased Visual Field Area With Age

Weale has suggested that changes in the ocular media, coupled with decreased pupil size are major contributors to the decrease in visual field size with age.[7] Later reports by Weale and others have suggested that retinal and post-retinal factors may play a more significant role.[8,9]

In the group of people reported above, there was a weak, but significant, association between pupil size and I-2 visual field area. The scatter plot for these data is shown in Fig. 6. Pupil diameter was measured to the nearest millimeter using the reticle in the observation system of the Goldmann perimeter immediately prior to field testing. The coefficient of correlation (r) for these data was + .56. Only 33% of the variation in field area was related to pupil size. A similar analysis for age and pupil size gave a similar correlation (r = 0.57). If a multiple regression analysis is done to
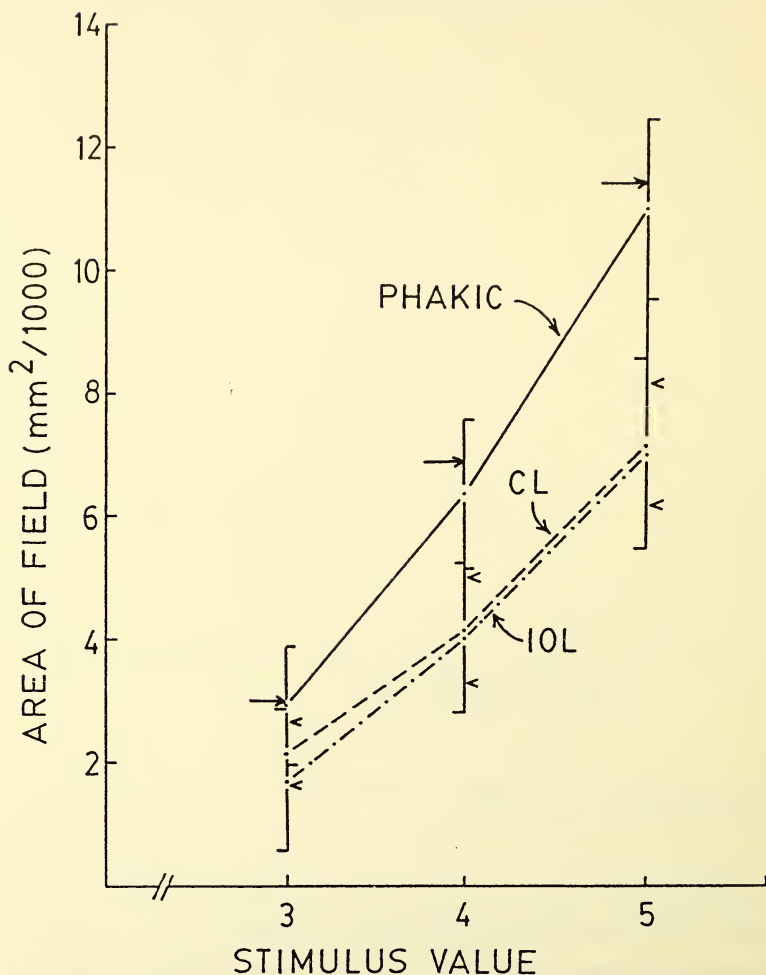
Figure 7. Visual field areas for two populations of post-cataract-surgery individuals (CL = corrected with contact lenses, and IOL = corrected with intraocular lens implant) and one population of phakic individuals. Age distributions for all three groups are comparable. Visual field area represented is that contained within the given isopter on the Goldmann field chart (see text for magnification compensation). Means and one standard deviation on either side of the means are shown. Reprinted with permission from Lazarus L, Williams TD. Visual field area in phakic, aphakic, and pseudophakic individuals. Am J Optom Physiol Optics 1988;65(7):593-7.

determine the combined effects of age and pupil size on field area, the correlation is only slightly increased ($r = 0.74$) from that between field area and age ($r = 0.72$).

To determine whether changes in the crystalline lens were contributing to the decrease in field area with age, a study was done on a group of pseudophakic individuals and a group of contact-lens-wearing aphakic individuals.[10] The logic behind the study was this: if the crystalline lens is responsible for decreasing visual field area in the older population, then an older population of pseudophakic or aphakic people should show a larger visual field than their age-matched phakic contemporaries.

Results of this study are shown in Fig. 7. The phakic population had a larger field under all test conditions. Statistically, there was no significant ($p < 0.05$) difference among the I-2 visual field areas of the three groups. For the I-3 and I-4 stimuli, however, the differences were significant at the 0.05 level. All of these analyses were done after correction for the effects of magnification. The right-facing arrows in Fig. 7 indicate a predicted field area for the average age of the phakic group (66.8 years). The average age of both the contact-lens- and intraocular-lens-corrected aphakic groups was 63.8.

These studies support the concept that pupil size and lens changes are not major contributors to the decrease in visual field size with age. The separation of the groups shown in Fig. 7 suggests that the population who develop cataracts may be fundamentally different from the age-matched population who do not. In Fig. 7, one notices that the area of the IOL and CL fields departs increasingly from the norm and from the phakic population as stimulus intensity is increased: this is similar to the behavior of the right visual field data shown in Fig. 4D. This suggests that the ganglion cells in the aphakic population may not be functioning normally. It may be that, for the IOL and CL populations depicted in Fig. 7, some common disturbance caused both the cataract and visual field changes.

BIBLIOGRAPHY

1. Drance SM, Berry V, Hughes A. 1967 The effects of age on the central isopter of the normal visual field. Can J Ophthalmol 2(2):79-82.

2. Hart WA, Becker B. 1977 Visual field changes in ocular hypertension: a computer-based analysis. Arch Ophthalmol 95(7):1176-79.

3. Williams TD. 1983 Computer-based analysis of visual fields: age-related norms for the central visual field. Can J Optom 45(4):166-170.

4. Williams TD. 1983 Aging and central visual field area. Am J Optom Physiol Optics 60(11):888-891.

5. Williams TD. 1986a Proportional increase in visual field size. Am J Optom Physiol Opt 63(9):782-3.

6. Williams TD. 1986b Visual field area response to increased target intensity: a method of detecting ocular disease. Am J Optom Physiol Opt 63(1):28-31.

7. Weale RA. 1963 The aging eye. New York: Harper & Row, p. 168.

8. Weale RA. 1982 Senile ocular changes, cell death, and vision. Pp. 161-171 in Sekuler, Kline, and Dismukes, eds., Aging and Human Visual Function, New York: Alan R. Liss.

9. Johnson CA, Adams AJ, Adams CW, Lewis RA. 1988 Evidence for a neural basis of age-related visual field changes. Pp. 44-47 in Noninvasive assessment of the visual system, 1988 Technical Digest Series Volume 3, Optical Society of America.

10. Lazarus L, Williams TD. 1988 Visual field area in phakic, aphakic, and pseudophakic individuals. Am J Optom Physiol Optics 65(7):593-7.

MEASUREMENT OF VISUAL ACUITY—TOWARDS STANDARDIZATION
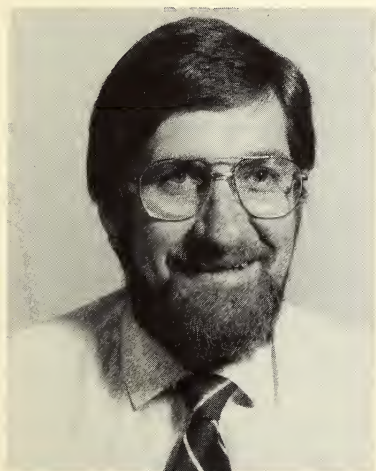
Ian L. Bailey
Professor of Optometry and Physiological Optics and Director of the Low Vision Clinic at University of California at Berkeley. He is a member of the National Academy of Sciences/National Research Council Committee on Vision.

ABSTRACT. Many variations of Snellen's original acuity test exist. This has led to confusion in recording and comparing visual acuities. Various attempts to standardize visual acuity notations are discussed. The author offers five recommendations that, if followed, would eliminate most of the remaining problems in achieving standardized visual acuity designations.

MEASUREMENT OF VISUAL ACUITY—
TOWARDS STANDARDIZATION

It is over a century since Snellen[1] published the description of his famous chart and introduced the Snellen Fraction that is still commonly used to designate visual acuity. Since then, there have been a multitude of variations of Snellen's original chart design and there have been several alternative methods proposed for expressing the measurement of visual acuity. While clinicians and visual scientists commonly use terms such as "Snellen acuity" and the "standard Snellen chart", there is a general ignorance of how nonstandard "common" testing procedures are. I know of no chart available today that uses either Snellen's original set of letters, his progression of size, or his chart layout. However there are numerous charts that manufacturers, suppliers, clinicians, and visual scientists might loosely refer to as "Snellen charts", but there is so much variation in significant design features that they can not be thought of as being standardized.

Despite a multitude of suggestions and effects to achieve some standardization[2] there has been little progress until recently. In the past decade, there have been serious attempts by three important authoritative bodies to develop better standardization of visual acuity measurement. Overall, there has been some encouraging but imperfect trends towards development of a standardized methodology for measuring visual acuity. The three major authoritative bodies that have considered standardization of visual acuity are: (a) the National Research Council[3], Committee on Vision, Working Group 39, 1979; (b) the Concilium Ophthalmologicum Universale[4], Vision Functions Committee, 1984; and (c) the International Standards Organization[5], draft proposals DP 8596 and DP 8597. Table 1 shows that there are several areas of general agreement between these bodies. There is agreement that a Landolt Ring target should be the standard optotype, alternative optotypes such as letters are generally encouraged; and what may be surprising to many is that the three organizations recommend a standard 4 meter test distance (this is in keeping with Hofstetter's[6] recommendation). All agree that the size progression should be in 0.1 log unit steps (size ratio 1.26 x). There is also agreement on the

methodology by which alternative systems (such as letters) should be calibrated against the standard Landolt Ring. The recommended procedure is to use a forcedchoice paradigm involving brief presentations of single optotypes. In my view, this calibration procedure is inappropriate because optotypes that yield identical visual acuity scores under conditions of single letter presentation need not necessarily (and in fact they do not) yield identical scores when the symbols are arranged in a chart format. I would strongly advocate that a standard Landolt Ring *chart* be specified and alternative optotypes or test devices should be calibrated against such a standard chart.

There are some clear points of disagreement between the three organizations' recommendations (see Table 2). The recommended number of letters at each size level is 10 for the NRC Committee on Vision, but the other two organizations have a requirement that there be at least five. A significant difference between the three sets of recommendations relate to the spacing between rows and between letters. The NRC Committee on Vision allows the spacing to vary between 1 or 2 letter heights at any level on the chart. The COU allows a choice of spacing ratio (and again this may be between 1 and 2 letter heights) but sensibly they insist that the spacing ratio be constant throughout the chart. The ISO recommends an irregular sequence of spacing ratios varying from 4/10 of a letter height at the largest size level (20/400) and 3 times the letter height at the smallest size level (20/16 and smaller).

Over the past ten years the National Eye Institute have funded several large multicenter research projects in which visual acuity was the critical outcome variable used to indicate the success of the medical intervention under consideration. For all such studies the similar visual acuity testing procedures have been used. The charts incorporate the design features that Jan Lovie Kitchen and I proposed in 1976[7]. The design features ensure that the task be standardized at each size level and this requires that there be the same number of letters per row, (we chose 5), that the spacing between adjacent letters and adjacent rows be proportional to the print size (spacing = 1 letter height), and that the size progression be logarithmic (constant ratio 1.256 x). Some of the NEI studies use charts that employ British letters[7] (5 units high, 4 units wide) while others use charts that have Sloan letters[8] (5 units high, 5 units wide).

Thomas Raasch and I have conducted several experiments to study factors that influence the magnitude and the reliability of scores of visual acuity. Here I will briefly present some of our findings that relate to (a) methods for designing visual acuity scores, (b) comparisons between different optotypes, (c) choices of spacing arrangements, and (d) choices of illumination levels. For these studies, our basic methodology has been to prepare printed charts that incorporate the required design features, and

TABLE 1

AGREEMENTS

	NRC WG 39	C.O.U.	I.S.O
Standard optotype	Landolt ring 4-choice	Landolt ring 4-choice	Landolt ring 8-choice
Alternative optotypes	Encouraged (Sloan O.K.)	Encouraged	Accepted
Test Distance	4 meters	4 meters	4 meters
Size progression	0.1 log unit (Ratio 1.2589x)	0.1 log unit (Ratio 1.2589x)	0.1 log unit (Ratio 1.2589x)
Calibration method	Single optotype protocol	Single optotype protocol	Single optotype protocol

TABLE 2

DISAGREEMENTS

	NRC WG 39	C.O.U.	I.S.O
# letters at each size	10 (2 gps of 5)	at least 5	at least 5 (5,8, or 10 OK)
Space between letters	1-2 letter hts may vary	1-2 letter widths uniform	0.4 letter ht at 6/120 1 letter ht at 6/95-6/48 1.5 letter ht at 6/38--6/19
Space between rows	1-2 letter hts may vary	1-2 letter widths uniform	2 letter ht at 6/15-6/6 3 letter ht at 6/4.8 -
Luminance	85 cd/m^2	at least 80 cd/m^2	120 cd/m^2 refraction 240 cd/m^2 daylight equiv

for each chart design we make two charts that are identical except in the order in which the letters were arranged. The two charts of a pair could be expected to yield the same score visual acuity since there is no significant design difference between them. Discrepancies between the two visual acuity scores from a chart pair becomes a metric for indicating the reliability, or repeatability, of the visual acuity test with that chart. Pooling discrepancy values across subjects allows for a broad cumulative measure of reliability. In assigning visual acuity scores, we use a system that assigns credit for every letter read. Visual acuity is expressed as logMAR and our charts have a size progression of 0.1 logMAR units per row. Since there are five letters per row, we assign 0.02 logMAR units per letter. Figure 1 shows a histogram of discrepancies collected in one of our experiments. This histogram is based on 189 comparisons. It may be seen that about 20% of the comparisons showed perfect concordance between the two scores of acuity from pairs of charts. In 33% of cases (16 + 17) there was only a one letter discrepancy between the two tests being compared, so that in 53% of cases the discrepancy was one letter or less. In 97% of cases the letter discrepancies were four letters or less so that only in 3% of cases did the discrepancy reach five letters (which is equivalent to one full row). The mean of this distribution reflects the average improvement or decrease in acuity on the second test. The mean is invariably close to zero. The standard deviation of the discrepancy distribution reflects the repeatability of the visual acuity score and the standard deviation of acuity scores will be 2 smaller than the standard deviation of the discrepancies.

Methods for Assigning Visual Acuity Scores

It is not an uncommon practice to assign visual acuity scores in increments that represent the range of sizes available. A standard criterion such as 60%, 70%, or 80% correct may be used to decide whether the subjects will be given credit for having read a given row satisfactorily. Such a method substantially impairs reliability. It is far better to give partial credit within a row, and this is most readily achieved by giving credit for every letter read. Table 3 presents analyses of scores that come from the data set represented in Figure 1. It shows the distribution of discrepancies (unsigned) when two different methods of visual acuity scoring have been applied. The first scoring system gives equal credit for each letter read (0.02 logMAR units per letter), while the other uses a criterion that the subject will be credited with having read a row successfully if 60% (3 of 5) of the letters are read correctly. There is perfect concordance between test and retest in 20% of cases for the first method, but for the second the concordance is perfect in 61% of cases. For the second method shows a discordance equal to one row of acuity in 39% of cases and a full two lines
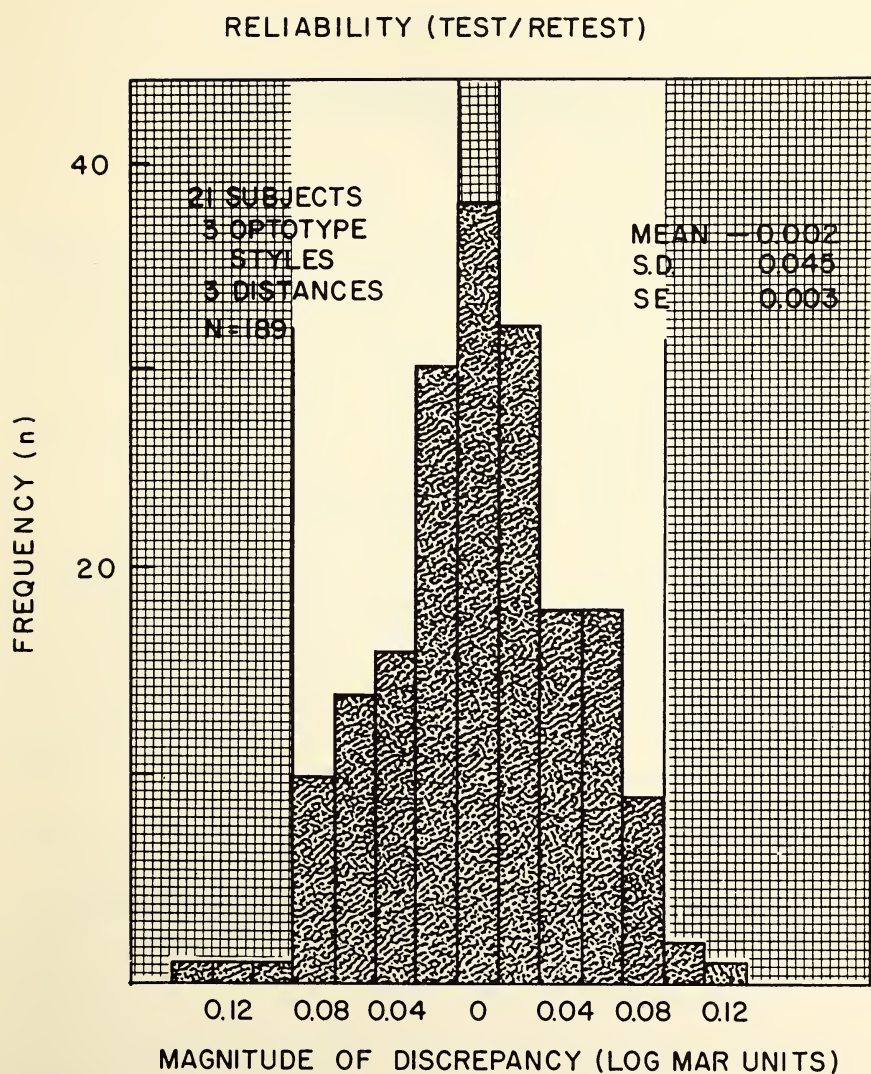
Figure 1

Distribution of discrepancies for test/retest visual acuity scores for 189 comparisons made with 21 subjects. Visual acuity scores determine by giving equal credit for every letter read. Shaded areas to the sides indicate when test/retest discrepancy equalled or exceeded one full line of acuity.

Ian L. Bailey

TABLE 3

Discrepancies between VA scores from pairs of Equivalent Charts
(Discrepancy = VA score #1 - VA score #2)

<u>DISCREPANCY</u>		<u>FREQUENCY (%)</u>			
# letters	LogMAR	<u>Equal credit</u> for each letter read		<u>VA criterion</u> - given row if 60% correct	
0	0.00	2 0	(2 0)	6 1	(6 1)
1	0.02	3 3	(5 3)	- -	- -
2	0.04	1 8	(7 1)	- -	- -
3	0.06	1 7	(8 8)	- -	- -
4	0.08	9	(9 7)	- -	- -
5	0.10	1.5	(9 8)	3 8	(9 9)
6	0.12	1	(9 9)	- -	- -
7	0.14	0.5	(1 0 0)	- -	- -
8	0.16	.	.	- -	- -
9	0.18	.	.	- -	- -
1 0	0.20	.	.	1	(1 0 0)

Data here from 189 comparisons.

Applying an <u>80% or 100%</u> correct <u>criterion</u> to same data set shows the same pattern.

Similar findings from other studies.

TABLE 4

Variations in visual acuity scores
(change expressed as % of one line)

	Repeatability (std dev)	Score rel to 4-choice Landolt	2-fold change in spacing	2-fold change in luminance
Landolt Ring 8-choice	41%	-37%	53%	
Landolt Ring 4-choice	36%	-----	66%	9%
British (5 x 4) letters	25%	13%	30%	
Sloan (5 x 5) letters	32%	43%	37%	15%

Experiments used normally sighted subjects,
Chart - geometric size progression (1.26X), 5 optotypes per row

discrepancy between scores from equivalent charts in 1% of cases (2 of 189). The method of giving equal credit for each letter only producing discrepancies that equal or exceed one row (5 letters or more) in 3% of cases. Clinicians or researchers interested in reliability of detecting changes of one full line of acuity, would be ill advised to use the method of assigning visual acuity scores by any methods that do not give credit for every letter read.

Visual Acuity Scores with Different Optotypes

Most of our investigations into parameters effecting visual aucity measurement have concentrated on four different optotype families, the four-choice and eightchoice versions of the Landolt Ring and the Sloan and British families of letters. These four different kinds of optotypes show differences in reliability, differences in the scores of visual acuity, and they show differences in their dependency on spacing and or on luminance. Table 4 shows some of the differences. In this table variations have been expressed as a percentage of one line of acuity where one line of acuity represents 0.1 logMAR units. It can be seen that Landolt Ring acuity scores are less repeatable than are letters obtained with the families of letters. The eightchoice version of the Landolt Ring shows least reliability. The eight-choice Landolt Ring yields a score that is 37% of a line poorer than the acuity from the fourchoice Landolt Ring. Relative to the fourchoice Landolt Ring, the British letters yield a score that is marginally superior, while the Sloan letters yield a visual acuity that is on average about 40% (2 letters)

TABLE 5

SPECIFICATION OF VISUAL ACUITY

Minimum Angle of Resolution (MAR)	0.5	1'	2'	5'	10'	20'
Decimal	2.0	1.0	0.5	0.2	0.1	0.05
Snellen (metric 6 m) (metric 4 m)	6/3 4/2	6/6 4/4	6/12 4/8	6/30 4/20	6/60 4/40	6/120 4/80
Snellen (feet)	20/10	20/20	20/40	20/100	20/200	20/400
Visual efficiency (%)	109%	100%	83.6%	48.9%	20%	3.3%
Log MAR	-0.3	0.0	0.3	0.7	1.0	1.3
Visual Acuity Rating (VAR)	115	100	85	65	50	35

of a line better. In our studies of spacing effects, we have considered spacing as narrow as 1/2 letter heights to as wide as 3 letter heights. Table 4 shows that Landolt Ring acuity is substantially more effected by spacing than are the acuities from letter charts. On the other hand, the Landolt Ring families are less effected by changes in luminance than are the acuities from the letter families. It should be noted that the changes in score due to relatively modest (two fold) changes in spacing can produce significant (1.5 to 3 letters difference) in visual acuity scores.

Methods for Designating Visual Acuity

There are many alternative methods for designating the value of visual acuity score. There is the familiar Snellen Fraction which in the United States is usually expressed for a standard test distance of 20 feet, and when the metric units are used it has been common to use 6 meters as the standard test distance. Currently however a test distance of 4 meters is now being advocated. Table 5 presents visual acuity scores using several methods of designation. In Europe decimal notation is popular and within the communities of clinical researchers and visual scientists it is common to express visual acuity in terms of Minimal Angular of Resolution or in terms of the logarithm of the Minimal Angle of Resolution. In my view, there is a need to change from the Snellen Fraction because it often leads to confusion whenever the test distance is other than common 20 feet (or 6 meters). There is a strange clinical practice of expressing the size of print in near vision tests in terms of an equivalent Snellen Fraction in which the numbers refer to measurements relating to distance visual acuity. An example of the potential for confusion is the following; an individual with a distance

visual acuity of 20/100 should be expected to achieve a "reduced Snellen of 20/50" if the testing is done at 20 cm, as it may well be. What then is this patient's near visual acuity? Another example of confusion resulting from misuse of the Snellen Fraction notation has arisen with the introduction of charts for which the intended working distance is 4 meters rather than the usual 6 meters (or 20 feet). On some of these new charts (Ferris et al[8], Strong and Woo[10]), the largest print on the chart subtends 5 minutes of arc at 40 meters and so an individual who can just read the top row from 4 meters should be awarded a visual acuity score of 4/40. However, the top line has been labeled 20/200. Confusion becomes manifest when the chart is presented at any distance other than 4 meters for which the chart was designed. For example, a patient reading the top (labeled 20/200) row when the chart is at 6 meters (say 20 feet) would actually have a visual acuity of 20/133. I believe the source of the problem in using Snellen Fractions to specify visual acuity is that the Snellen Fraction specifies two quantities, the first for test distance and the second is a measure of print size, and together these represent an angle. It is the angle, a single dimension, that represents the acuity.

Here I wish to propose a new system for rating visual acuity. I am suggesting a 100 point scale on which visual acuity of 20/20 is represented by 100 points, 20/200 as 50 points, and 20/2000 as 0 points. This Visual Acuity Rating scale follows a logarithmic size progression (VAR = 100 50 logMAR). The scaling is such that 15 points on the scale represents a doubling of visual angle; 20/20 is 100 points; 20/10 is 115 points; and 20/40 is 85 points and so on (see Table 6). On this scale 5 points represents one line of acuity where a 0.1 logMAR size progression has been used. Should there be 5 letters per row, each letter carries a weight of 1 point. On charts that follow the Bailey Lovie design principles the task of scoring visual acuity becomes one of counting letters. Should a nonstandard test distance be used (for instance the chart moved from the standard test distance to a closer distance), an appropriate adjustment can be made to the test scores simply be subtracting a constant value. For example, if the chart were to be moved to half the standard viewing distance, the patient would be expected to read 15 more letters. Then the visual acuity rating score would need to be adjusted (by 15 points) in accordance with this closer viewing distance. There are many features of this scale that cause it to feel about right. The 50 point level is the traditional cutoff for "legal blindness", the 25 point level is equivalent to 20/63 and this is a common cutoff for eligibility for low vision services for those who are not legally blind. Furthermore, the 25 point level (this is equivalent to 20/630) is about the point at which it becomes very difficult to achieve significant enhancement of visual performance through the use of optical low vision aids. In some ways, this scale is similar to the visual efficiency scale proposed by Snell and Sterling

TABLE 6

"Visual Acuity Rating"

VAR = 100 - 50 (LogMAR)

5 rating points per row if size progression ratio =1.26 x

115	20/10	6/3				
110	20/12.5	6/3.8				
105	20/16	6/4.8				
100	**20/20**	**6/6**		**50**	**20/200**	**6/60**
95	20/25	6/7.5		45	20/250	6/75
90	20/32	6/9.5		40	20/320	6/95
85	20/40	6/12		35	20/400	6/120
80	20/50	6/15		30	20/500	6/150
75	*20/63*	*6/19*		25	*20/630*	*6/190*
70	20/80	6/24		20	20/800	6/240
65	20/100	6/30		15	20/1000	6/300
60	20/125	6/38		10	20/1250	6/380
55	20/160	6/48		5	20/1600	6/480
50	**20/200**	**6/60**		**0**	**20/2000**	**6/600**

If 5 letters per row, then 1 point per letter

Change viewing distance, then change assigned score by a constant

in 1929. Their scale was introduced to provide a yardstick for gauging compensation for vision loss. They had assigned a visual efficiency of 100% for a visual acuity of 20/20 and 20% for 20/200. The SnellSterling visual efficiency scale seems to work well in the range 20/40 20/160 region of the acuity spectrum, but at the extremes equal steps on the scale seem to represent very uneven differences in visual performance. I would argue that, on the scale I am proposing here, relatively equal decrements in visual performance are indicated by equal number of units of change on the VAR scale. This would be expected from Westheimer's observations and recommendations[11] regarding the appropriate scaling of visual acuity. The VAR scale is most easily applied to charts of the BaileyLovie design (Fig 2) but it can readily be used for more traditional chart layouts. Figure 3 illustrates a common chart layout on a "Standard Snellen Chart" but VAR
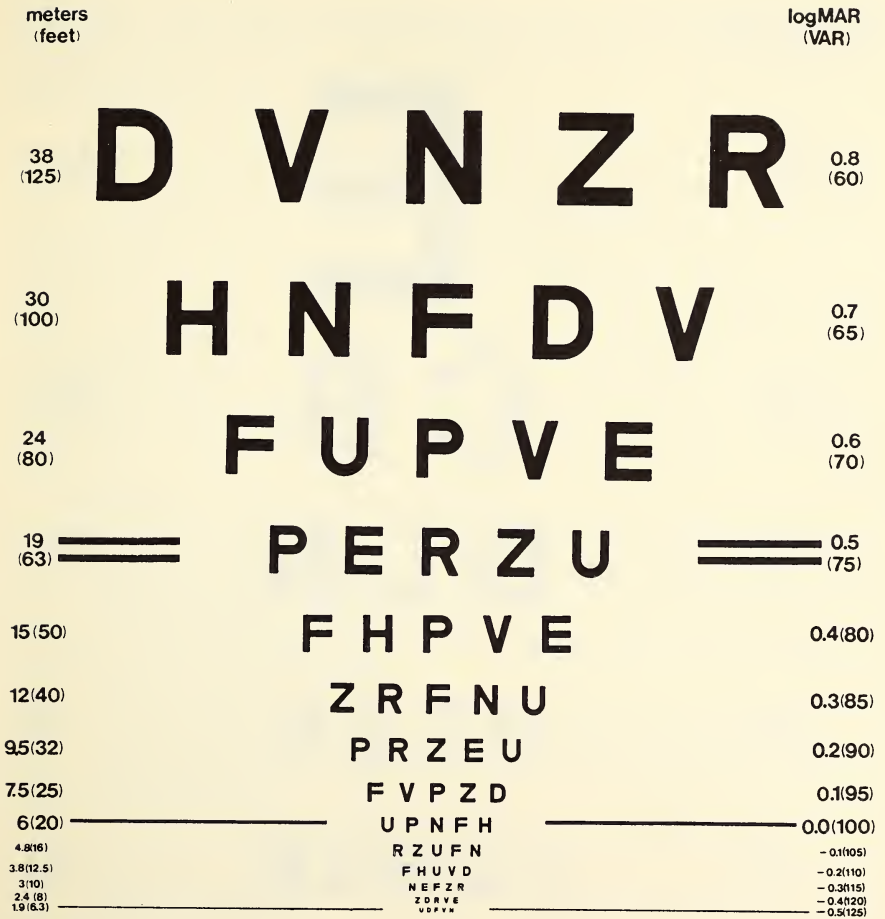
Figure 2
A Bailey-Lovie chart that uses British (4x5) letters and includes VAR scaling.

Ian L. Bailey

Snellen

VAR

Snellen		VAR
200	**E**	50
100	**C B**	65
80	**D L N**	70
60	**P T E R**	76
50	**F Z B D E**	80
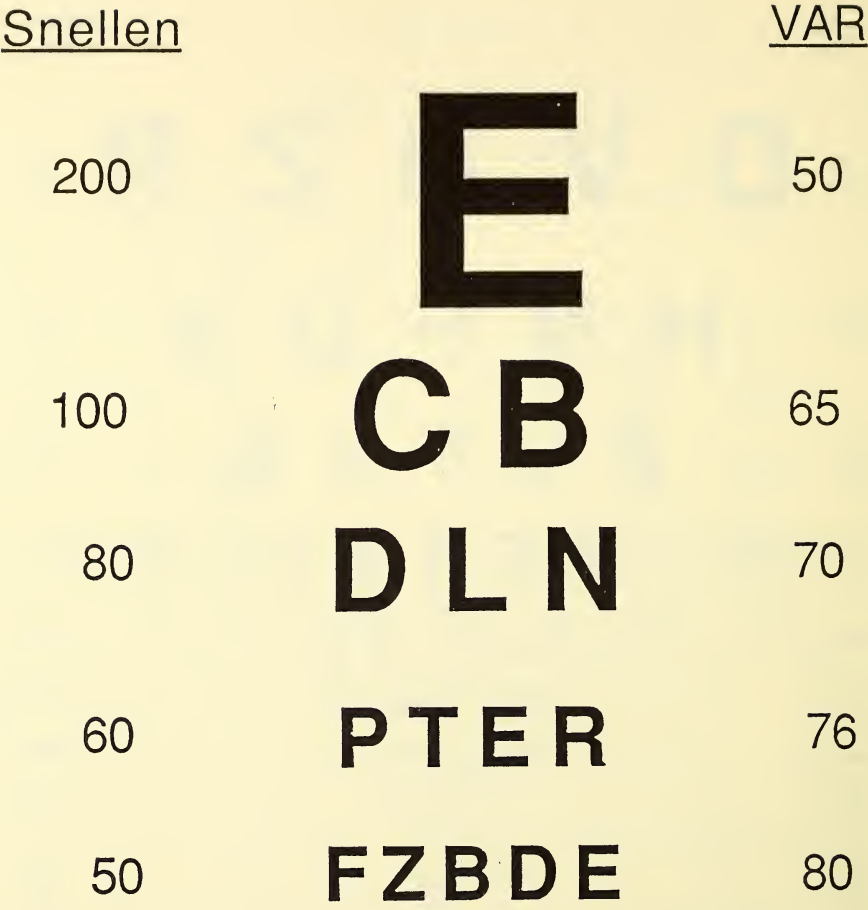
Figure 3
  Representation of a section of a common "Snellen chart" layout and here VAR scaling
  has been added.

scaling has been added. Interpolation can easily be used; for example, a patient who only reads the large E and, let us say, the C on the next row would receive a VAR score of 58. Scaling in VAR unit can be applied to any visual acuity test.

It should be recognized that clinicians are likely to be reluctant to change their ways and adopt a new scoring system for visual acuity measurement. However given that there is currently significant confusion that relates to the clinician's sloppy use of Snellen Fractions, I feel it is justified to recommend a system that uses a single number to indicate the angular dimension in which visual acuity should be expressed. The VAR scale is one on which equal points carry about equal value in terms of visual performance or efficiency, and furthermore it is a scale that may facilitate communicate between clinician and patient. For instance, it is relatively easy to tell a patient that visual acuity is a little better than normal as it is 105 points, where normal vision is considered to be 20/20 or 100 points. Clinicians and patients can easily identify with these values.

In summary I would advocate the following:
(1) That the four-choice Landolt Ring target be the standard optotype and the standard test of visual acuity be measurements of visual acuity made with a Landolt Ring chart designed so that it complies with the design principles of Bailey and Lovie.
(2) Letters charts or charts of alternative optotypes should be calibrated against a Landolt Ring chart.
(3) The spacing between adjacent optotypes on any one row should be equal to the width of those optotypes and the spacing between adjacent rows should be equal to the height the optotypes of the smaller of the two rows.
(4) Visual acuity scores should be assigned using a method that gives credit for every letter read correctly.
(5) That a new scaling system Visual Acuity Rating be introduced to specify visual acuity.

References
1. Snellen H. Letterproven tot Bepaling der Gezigtsscherpte, Utrecht, Van Der Weijer, 1862.
2. Bennett A.G. Ophthalmic test type. Brit. J. Physiol. Optics 22, 238-271, 1965.
3. N.R.C. Committee on Vision. Recommended standard procedures for the clinical measurement and specification of visual acuity 1979. Advances Ophthal 1980, 41, 103-148.
4. Concilum Ophthalmologican Universale, Visual acuity measurment standards. COU Vision Functions Committee, 1984.
5. International Standard Organization, DP8596 and DP8597.
6. Hofstetter H.W. From 20/20 to 6/6 or 4/4. Am J Optom, 1973, 50, 212-222.

7.   Bailey I.L. and Lovie J.E. New design principles for visual acuity letter charts. Amer J Optom and Physiol Optics 53, 740-745, 1976.

8.   Ferris F.L., Kassoff A., Bresnick G.H., and Bailey I.L. New visual acuity charts for clinical research. Am J Ophthal 1982, 94, 91-96.

9.   Bailey I.L. and Raasch T.W. Standards for the measurement of visual acuity. Amer J Optom and Physiol Optics 1984, 61, 119P-120P.

10. Strong G. and Woo G.C. Distant visual acuity chart incorporating some new design features. Arch Ophthal 1985, 103, 44.

11. Westheimer G. Scaling of visual acuity measurement. Arch Ophthal 1979, 97, 327-330.

A MODEL OF VISUAL REHABILITATION WITH SPECIALIZED OPTICAL TECHNOLOGY
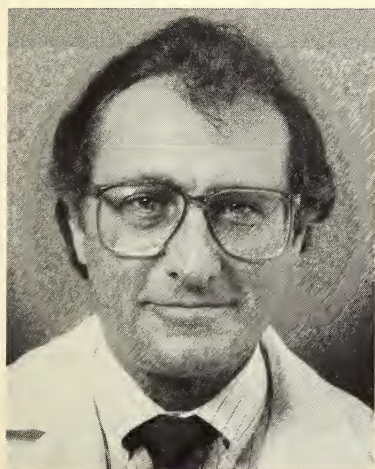
Joseph R. Zahn
Director of Low Vision Services, Baptist Eye Institute, Knoxville, Tennessee.

ABSTRACT. Improved technology permits more persons suffering visual disabilities to benefit from special optical aids. At the same time the number of persons needing vision rehabilitation is increasing as the elderly population increases. The full range of vision rehabilitative services must be available at a given site if optimum rehabilitative care is to be provided and if adequate time is to be given to each patient. This paper describes an interdisciplinary practice model designed to ensure the various competencies required, provide essential instrumentation, and maximize cost effectiveness.

A MODEL OF VISUAL REHABILITATION WITH SPECIALIZED OPTICAL TECHNOLOGY

The number of individuals who require special optical aids to read newsprint is growing significantly in most countries of the Western World[1], due to untreatable eye disease to decreased visual function after medical or surgical management[2,3]. Age related maculopathy following successful cataract extraction is but one example.

Those who can read newsprint only after specialized optical intervention have been identified by authors as a unique disability group requiring a variety of treatment programs[4-9]. Symposia describing these programs and urging development of special centers to serve this population have been sponsored by several organizations concerned with eyecare and visual rehabilitation. In spite of this interest there has been very little progress in providing added visual rehabilitative services for this large and growing group since it was distinguished from other groups suffering visual disability more than 20 years ago[9,10].

According to the American Foundation for the Blind there were 197 locations in the United states in 1984 that provided rehabilitative services for those who are partially sighted. This represents modest growth and is far short of the number needed to provide vision rehabilitative services. Prevalence data concerning vision disorders that lead to low vision in most instances can only be estimated. One estimate suggests that as many as two million people in the United States suffer from age related maculopathy, one of the most common of many ocular diseases that lead to low vision. If all are to receive attention, an average of approximately 10,000 patients would be seen at each of the service centers. It is estimated that each eye care specialist could manage only about 16 new patients each week. If this patient volume represented the workload of one eye care professional at each center, then by this accounting it would require approximately 12 years to serve those presently afflicted by age related maculopathy alone.

Services to those who might be rehabilitated so that they can read newsprint is limited by many factors. Early in this century, when vision rehabilitation was just beginning, distinctions were not made on the basis

of what vision performances might be achieved. Sighted patients were often taught alternative sensory input by techniques now reserved for the totally blind. Later classifications based on legal definitions of blindness or on crude tests such as finger counting provide little predictive value for visual function; nor do various commonly used instrumentation procedures such as fundus photography or visual fields. While the optical principles of magnification are straight forward, various reading tasks performed under magnification require special adaptations. Observing the detail on a stamp, for example, is a much simpler procedure than locating points on a world globe. The practice of referring patients to a drugstore for magnification devices did not take into account these differences, and is responsible for a sense of futility among many patients and practitioners.

For these reasons and because of the considerable time involved in patient orientation and training, as well as testing, low vision patients tend not to receive optimum care outside these special centers. The experience of an opththalmology resident who was asked at the end of his residency what he had learned from this year in a low vision clinic is illustrative. He said, "I've learned that I will refer my patients with decreased visual function not corrected—or only partially corrected by medical and surgical means—to a low vision clinic." The author believes this is the appropriate response of family practitioners of eyecare and that it will become the standard for most practitioners. Therefore, it supports added development of low vision clinics within departments of ophthalmology, schools of optometry, and within vision foundations and agencies. Provision of services to the low vision patient is very time consuming, requires the competence of individuals with various kinds of special training, and requires a high degree of patience on the part of practitioner and patient. For these reasons and because it is not a cost efficient activity as currently formed, it is important that one or more new models be developed. A description of one model that is in evolution follows[14-17].

An Interdisciplinary Model of Visual Rehabilitation

Granting that there are a number of different types of private practice arrangements, this model assumes one in which there exists a group of individuals, each of whom independently operates a practice. It proposes for low vision an interdisciplinary team with three full time staff members (director/evaluator/researcher; coordinator/educator; receptionist/secretary). This model contends, unlike the multidisciplinary models,[3-6] that an individual can "wear more than one hat". Each member is multidisciplinary; the team is interdisciplinary. The duties of each member are shown.

DIRECTOR/EVALUATOR/RESEARCHER

DIRECTOR	**EVALUATOR**	**RESEARCHER**
Funding	Testing	Basic
Grants	Training	
Fees	Lending	Applied
Donations	Referral	
Public Education		Publications
Advocacy		

Director/Evaluator/Researcher

The director provides advocacy, public education, and funding. These are accomplished through support groups, lectures, speeches to organizations, and both grants and fee for service. The evaluation includes testing for visual function, identification of useful optical and alternative devices, and initiation of a training program. Research can be long term studies, case reports, and the guidance of those with interships.

COORDINATOR/EDUCATOR

COORDINATOR	**EDUCATOR**
Scheduling	Intake screening
Equipment	Training
Personnel	Follow-up

Coordinator/Educator

The coordinator/educator provides screening, intake, training and follow-up. The screening test permits an analysis of useful visual function of those in schools, nursing homes, and, it is a tool to demonstrate to some who may not know about their visual function a knowledge of it. The intake nominally includes the history (medical, visual, career, social), the psychological and intellectual status, the support system (community, family), the utilization of eyeglasses and/or specialized optical and alternative devices, and the interchange of information among referral sources (medical, educational, vocational, residential).

Under the direction of the coordinator are the volunteers and the optician. The sale of optical devices is separate from the rehabilitative program.

RECEPTIONIST/SECRETARY

RECEPTIONIST	**SECRETARY**
Schedule	Correspondence
Brochure	Publication
Information	Grants
Referral	Accounts
	Orders

The duties include those listed but may not be limited to them. Other personnel may include a rehabilitation counselor and an optician.

Receptionist/Secretary

An underrated staff member is the receptionist/secretary. This individual provides information about the clinic, the staff, and the problems of partial sight. Thus in addition to the standard requirements of the "title", this individual addresses specific patient inquiries, designs and mails information packets, and submits forms for and from patients.

Application Of The Model

This model has taken five years to develop. Its success where success is defined as a full-time (6 days a week) occupation for each staff member, and fully funded by fee for service, was dependent upon funding resources, the referral base, and community support.

Figures 1 and 2 show data for a three year period for a low vision clinic located within the same city but at two geographical locations. The initial visits were for visual rehabilitation with specialized optical and alternative technology. They were not visits for surgery, medication, or standard refractions.

Three factors influenced 1983, the most erratic development year. Vacation for the staff influenced the July initial visits; thus, the follow-up visits. Parenthetically, our procedure requires the patient to return within one week after the initial visit (evaluation). At this time he has received a device and provided preliminary instruction.[7] The follow-up data are contaminated by patients who returned for further training and devices, sometime during the first year of rehabilitation.
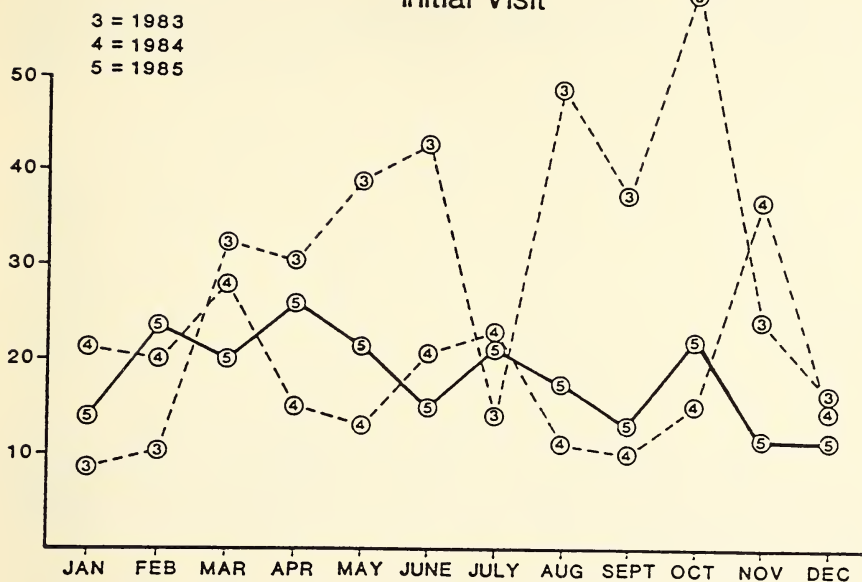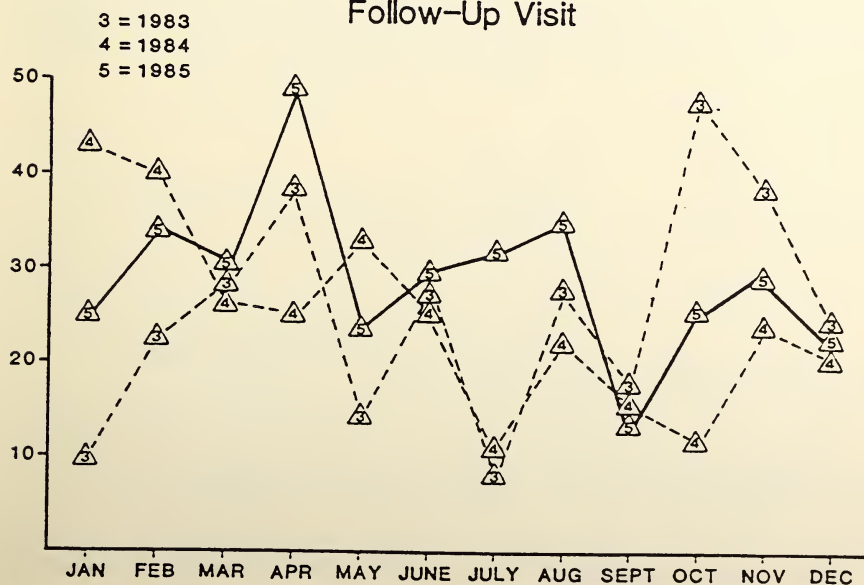
Figure 1

Initial Visit

3 = 1983
4 = 1984
5 = 1985

Figure 2

Follow–Up Visit

3 = 1983
4 = 1984
5 = 1985

A second factor was the referral sources. Retinal degenerations represent the largest referred group Based on a sample of 918 patients, 8% were referred by non-profit organizations (American Council of Blind, Lions, Odd Fellows and Rebekahs), 1% by agencies serving the blind, 19% from information obtained in television interviews and newspaper articles, 39% from information obtained from previous patients, and, 33% directly from eye care practitioners. The practitioners were mostly ophthalmologists (92%).

Success of the patient was not influenced by the source of referral. We identified three referral groups: Group I (eye care specialists), Group II (Agencies for the Visually Impaired, Department of Education), Group III (nursing homes, service organizations, self referral). The success for the patients that is, their continued use of the devices and techniques after a six-month follow up survey, showed Group I (79%), Group II (87%), and Group III (71%). There were 36% using one optical device, 22% using two, and 9% using three or more. A more thorough analysis is provided elsewhere.[16]

The third identifiable factor influencing 1983 was publicity. In May and September, the low vision clinic was spotlighted by the verbal and written media. Referral, telephone inquiries, and self-made appointments were significantly increased.

Figure 3 Survey form sent to 790 patients

Your opinion of your experience at the low vision clinic is valuable to us. With your help we can improve our service to other patients. We hope you will take a moment to answer the following questions and return them to us in the enclosed stamped, addressed envelope.

We appreciate the opportunity to work with you at the low vision clinic.

1. Did you complete your low vision evaluation?
   _____ Yes   _____ No

2. Did you receive an optical aid (magnifier, glasses, CCTV, etc.)?
   _____ Yes   _____ No

3. Are you using your optical aid?
   _____ Yes   _____ No

4. If you are not using your optical aid, do you want us to contact you about it?
   _____ Yes   _____ No

   If yes, print name and telephone number: _____

   _____

Your Comments: _____

   _____

Previous data were obtained from the third data collection procedure, a survey form (Figure 3) sent to 790 patients (with 395 respondents, 310 of which were analyzed because 48 forms were blank, and 27 were uninterpretable due to double answers on questions 1 and 2). Selected survey results are shown in Table 1.

Table 1

Use of optical device	Request for Follow-up		
	Yes	**No**	**No Reply**
Yes	8	53	98
No	23	64	16
No Reply	3	22	23

Response information is extended by combining answers. Question 3, "Are you using your optical aid?", was matched with question 4, "If you are not using your optical aid, do you want us to contact you?" If the patient replied, "Yes.", that they were using their optical aid, they could have replied "Yes" (8), "No" (53), or "No Reply" (98) to our query for a follow-up. 150 of 310 (51%) were using their device, and 103 (33%) were not. We were uncertain about how the remaining 16% should be classified. The variance among the 125 who responded "No" or "No Reply" to the follow-up and who were not using their optical sevices should be explained. Of these, 74 (59%) did not receive an optical device. Of those who did receive an optical device, 48% needed a follow-up. These patients were contacted by telephone in April and August. They are successfully utilizing their optical devices following subsequent training.

Discussion

Visual rehabilitation with specialized optical devices and techniques is a needed service. Although numerous articles, symposia and seminars emphasize this as an effective approach to visual disability, there has been little application within a private practice. In this article an analysis of trends within the initial and follow-up visits, and a survey were used to determine the parameters of a successful low vision service.

The model presented here requires three full time staff members, two 10 x 12 rooms, and approximately $30,000 in testing, training, and lending equipment. Each staff member's duties were outlined. Each staff member is multi-disciplinary but functions within an interdisciplinary paradigm.

During the anlysis of the initial visits it was found that eye care specialists tend to refer patients because of their visual acuity level or an ad hoc motivational assessment. These criteria do not predict patient success. Referrals are influenced by publicity from television, newspaper, radio,

and previous patients. The manner in which the referral sources handle this state-of-the-art information influences the patient's success.

Follow-up visits are influenced by the patient's successful use of an optical or alternaitve device, and by the contact generated by the low vision service staff. Rehabilitation is difficult. The devices can be expensive and their cosmetic appearance displeasing. These factors influence follow-up. However if the patient completes our program, 87% are visually independent. If they complete the evaluation and receive devices but do not complete training (educational process) or follow-up procedures, 52% are successful.

These data permit certain generalizations to be made, and pertinent questions to be asked by private practitioners, department chairman, and financial officer.

1. It is advisable to investigate the community resources prior to investment in staff and space.

   Aside from your own clinic are there others in the community who support visual rehabilitation? Eye care practitioners, service for the visually impaired, and organizations such as the Lions, can be a beneficial referral base. Are there advocacy groups? Are there qualified rehabilitation teachers or orientation and mobility instructors in the community? Are there any strong preconceived bias among these individuals and groups? The data here suggest that any collaborative and educational activities made prior to clinic construction will significantly influence success.

2. It is advisable to investigate funding resources for the clinic and the patient.

   Can the patient/clinic get reimbursed for specialized optical and alternative devices and techniques, and, for the evaluation. Third party carriers including Medicare, do not usually fund the evaluation or devices. This has a number of implications. For the billing service it means that it is not cost effective to submit insurance forms. For the low vision service it means that only those with adequate personal funds will seek the service. For the society it means that many individuals will remain visually dependent. And it has meant that eye care practitioners, sincere in their effort to assist the partially sighted, will bill for a standard eye examination rather that a low vision evaluation. Since one trend in our data indicated that those who require the "strongest" available devices are referred initially, it is important that equipment resources be directed here. Since there is little third party funding, the service managers shouyld vigorously seek organizational, foundation, and local community funds for equipment, both testing and lending.

3. It is advisable to consider the patients as multiply handicapped, thus influenced by environmental and medical factors.

   A service should have adequate parking facilities with direct access, and more important a van to deliver patients. Check the public trans-

portation system to determine accessibility and scheduling. Do nursing homes and schools provide transportation? Except fewer patient visits in the winter. Develop a relationship with those working with diabetes, stroke victims, neurologically impaired, and multiple sclerosis. There is an expanding category of those with central vision loss, who will require visual rehabilitation.

To develop a low vision service requires preplanning and collaboration. Low vision services will not be sustained successfully if they operate under the assumptions that low vision care involves only prescribing magnifiers, that non professional staff can evaluate patient requirements, or that lending and training are not necessary. In this and other articles the author has attempted to describe a model that avoids these pitfalls and ensures successful visual rehabilitation of patients with central and peripheral visual impairment by considering all necessary parameters of diagnosis and treatment.

References

1. Kerchner, C. and Petersen, R. (1979) The latest data on visual disability from NCHS.J. Vis. Imp. and Blindness 73(40), 151-153.
2. Lovie-Kitchin, J.P. and Bowman, K.J. (1985) Senile Macular Degeneration: Management and Rehabilitation.
3. Faye, E.E. (1976) Clinical Low Vision. Little, Brown and Co.
4. Apple, M.M., Apple, B.B. and Blasch, D. (1980) Low Vision. In Welsh, R.L. and Blasch, B.B. (Eds.) Foundations of Orientation and Mobility, Chap. 7, pp. 198-224. American Foundation for the Blind, New York.
5. DeStefano, A.F. (1984) Low Vision Research and Training Center, Pennsylvania College of Optometry. Yearbook of the assoc. of Educ. and Rehab. of the Blind and Visually Impaired, Washington, D.C.
6. Jose, R. (Ed.) (1983) Understanding Low Vision Am. Foundation for the Blind, New York.
7. Zahn, J.R. (1982) What can we do for the partially sighted? Nebraska Med. Journal 67(10), 281-285.
8. Lindquist, B., Bachman, O., Inde, K. and Trowald, N. (1978) Low Vision Rehabilitation. Uppasala, Sweden.
9. World Health Organization. Meeting: The use of Residual Vision by Visual Impaired Persons. Brussels, pp. 27, January 28-30, 1881.
10. American Foundation for the Blind (1984) Directory of Agencies Serving the Blind and Visually Handicapped in the United States, 22nd Ed., Sect. 2, 4, pp. 53-334.
11. Kohler, I. (1962) Experiments with goggles. Scientific Amer. 206, 62-72.
12. Riesen, A. (1961) Studying perceptual development using the technique of sensory deprivation. J. Nervous and Neurol. Disorders 132(1), 21-25.
13. Byer, A. (1979) Magnification: The goal of low vision lenses. Rev. of Optom. 3 47-48.

14. Zahn, J.R., Favero, B. and Horgan, J. (1985) Thoughts on Visual Acuity and Involutional Macular Degeneration. J. Vis. Rehab. 3(2), 10-11.

15. Zahn, J.R., Favero, B. and Horgan, J. (1988) Model of Visual Rehabilitation with Specialized Optical Technology. J. Vis. Rehab. 2, 1, 59-76.

16. Zahn, J. R., (1987) Predicting the Success of a Low Vision Clinic. J. Vis. Rehab. 1, 1, 35-39.

17. Zahn, J. R., (1988) Age-Related Maculopathy (ARM): The Rehabilitative Process. Wellness Perspectives 5, 1, 23-27.

TECHNOLOGY FOR VISUALLY IMPAIRED PERSONS: OPTIONS IN ORIENTATION AND MOBILITY
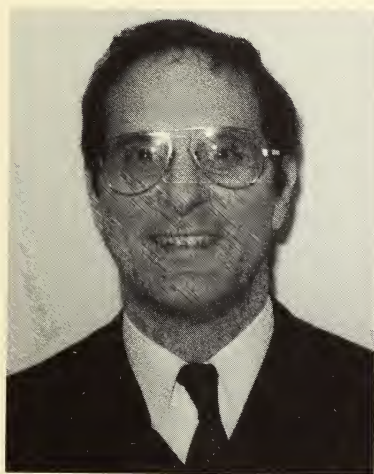
Mark M. Uslan
National Consultant in Orientation and Mobility at the American Foundation for the Blind, New York, New York. He is co-author of a book in press: *"The Profession of Orientation and Mobility"*.

ABSTRACT. The extended cane remains the principally used mobility device for blind persons. Its chief limitation is that it probes only the near space. This limitation is overcome by guide dogs; but here there is little sensory input, and few blind persons make this choice. The author describes new mobility devices that are designed to combine the high sensory input feature of the extended cane and the extended space feature of the guide dog. He recommends greater cooperation between scientists, engineers, practitioners, and instructors in order to address the remaining challenges.

TECHNOLOGY FOR VISUALLY IMPAIRED PERSONS: OPTIONS IN ORIENTATION AND MOBILITY

We often take for granted what it means to be able to move about at will and in safety — to be able to shop, meet friends and family, go to recreational events, and of course, commute to work and once there be able to do our jobs. This freedom of mobility is necessary to be able to lead a life of independence and dignity; without it a person is isolated, cut off from the mainstream of society and totally dependent on others. For most blind people the freedom of mobility is never taken for granted. Achieving it is the single most difficult obstacle they face.

While mobility assistance from friends and family members is convenient and unavoidable, true independence requires self reliance. When traveling by themselves, the vast majority of independent blind people depend upon a long cane. A small minority use a guide dog.

There is much significance in the fact that the preferred mobility device for blind people is nothing more sophisticated than a long cane. Despite the rapid pace of technological innovation today and despite the tremendous impact that this technological revolution has had on the lives of every member of the American society, the long cane has endured as the most appropriate assistive device for blind people. And it would not be surprising if the long cane continued to be the most used mobility device for blind persons through the next decade and perhaps even well into the 21st century.

Why is this the case? The cane is a primitive tool that only provides a limited solution to the mobility problem. Along with it, the blind traveller must endure the negative and erroneous stereotype of the long white cane as a tool used by those who are utterly dependent — the beggar's device. So why has it not been replaced by a more advanced technology, one that conveys less negative symbolism? First, with all its disadvantages, the cane is reasonably effective at scanning the path of travel and it does so at a reasonable cost and with a minimum of inconvenience. Second, the options are limited to only three other alternatives: the guide dog, optical devices for persons with sufficient usable vision, and expensive but questionably

useful electronic aids which are used as secondary devices to the cane or dog.

Oddly enough, the cane was used by blind people since before biblical times but it was only in the mid 20th century that its most useful attribute, its length, was recognized. For many years, and even today, blind people were given a short support cane which is virtually useless since it is not long enough to effectively scan the path of travel.

One of the more obvious shortcomings of the long cane's simple system, that it leaves the upper body vulnerable to low overhangs, does not always detract from its effectiveness, for such overhangs rarely obstruct pedestrian walkways. And the cane is more than a bumper to guard against collisions and a drop-off detector to prevent falls. The cane detects objects and surface changes and gives subtle vibro - tactile and acoustic feedback that can provide critically important clues to orientation. There are many examples: detecting the textural difference between a concrete sidewalk and the macadam of a driveway; detecting acoustic change as the boundaries of a path changes from a wall to a fence, to an open space; etc..

Once the long cane was recognized as an important mobility device, it was not long before a very useful design feature was added to it: the ability to be mechanically folded and put in one's pocket or pocketbook, and then taken out and unfolded at will. This folding feature made the cane especially attractive to persons with useful remaining vision as it enabled them to store the cane when they did not need it. The label of blindness which the cane announces in so visible a way, could be limited to those times when the individual really only functions as a blind person.

The major limitation of the cane as a travel tool is that it only probes space near to the traveller. Its effective range is limited to its extended length which is never more than about three feet. When used properly the cane can not be expected to do much more than address near space orientation and safety concerns so that the traveller can concentrate on the larger problem at hand: where he is in relation to where he needs to go. Nevertheless, by teaching cane travel techniques and the orientation skills associated with it, instructors of orientation and mobility (O&M) have made great strides in remediating what might be called the generalized disorientation problems of blind people.

The guide dog is even more efficient than the cane in freeing the traveller from the concerns of near space. Given the "forward" command, the dog will assist in keeping the traveller on the path and will steer clear of obstacles, enabling travel to proceed at a brisk pace, uninterrupted by contact with the environment. Unlike the cane however, the dog provides little sensory input —the traveller must have the desire, the self assurance, and the ability to make orientation decisions quickly and with less information of the type provided by the cane's continuous contact with the

environment. This is one of the reasons why the dog is preferred by only a small minority of blind persons. There are other reasons having more to do with lifestyle and convenience than ability, but to a large degree the success of the guide dog movement is based on a realistic trainee screening policy — only those blind persons who have the interest and the potential to handle the daily rigors of the guide dog obtain the training.

What about high tech solutions to the mobility problem? Of the five major electronic travel aids (ETAs) presently on the market, four provide ultrasonic sensing of objects in or near the path of travel. These devices offer options for hand held, chest mounted, cane mounted and spectacle mounted operation. The Laser Cane provides cane mounted optical sensing of objects and drop-offs. All together it has been approximated that 3,500 ETA devices have been produced since the 1960's[1] and that less than 700 people are presently using them even though in 1980 alone, some 20,000 blind individuals received O&M training.[2] In a recent report on ETAs the National Research Council[3] (1986:4) summarized the problem:

> "....Since the 1960's, many have expected that by exploiting the possibil-ities offered by a burgeoning electronics technology, it should be possible to build an electronic travel aid (ETA) that could provide much of the critical information about space that accounts for the ease of mobility of sighted pedestrians and that is not available to blind pedestrians. However, as the years passed and attempt followed attempt, it became apparent that the potential offered by ETAs was not being fulfilled. Although some improvement in performance could often be attributed to their use, in most cases the improvement was modest (Shingledecker and Foulke, 1978; Brabyn, 1982). The mobility of a blind pedestrian using an ETA should, in the opinion of some, approach the mobility of a sighted pedestrian (Leonard, 1968). Though many would regard this criterion as unrealistic, they would at least insist that the improvement enabled by an effective ETA should be more than modest. The ETA should make a difference that is obvious and indisputably significant. A successful ETA should enable independent, efficient, effective, and safe travel in unfamiliar surroundings.
>
> To date, no ETA has been built that can meet even the less demanding of the criteria just mentioned, and it is commonly recommended that an ETA be regarded as only an ancillary aid to supplement a primary aid: the long cane or the dog guide."

The Market for New Mobility Devices

There are a number of demographic factors that should be taken into account when targeting the market for any new mobility device. As a reference, the Social Research Department of the American Foundation for

the Blind has estimated that in 1987 there were almost 600,000 legally blind persons in the U.S..[4] In 1978 it was estimated that approximately 20% of legally blind persons were either totally blind or could only perceive light.[5] While we can assume that all of these blind people need a mobility device, that is not the case for all of the remaining 80% of the legally blind who have usable vision.

Data on O&M services in 1983 indicated that of those persons who received the services of O&M specialists, approximately 70% had better vision than light perception. Data on how many of these individuals needed a mobility device is not available but it can be assumed that some did not — O&M training without a cane and even without an optical device may have been sufficient for a number of these people.

There appear to be four distinct markets for mobility devices within the legally blind population. First, there are the non-visual travellers who use a cane or dog. It can be assumed that all of these people would benefit from a mobility device that would be a marked improvement over the cane or dog. Second, there are the non-visual travellers who do not use a cane or dog because of other impairments and are either institutionalized or homebound and totally dependent on care-givers. While these people would also benefit from a mobility device, the nature of that device and how it would be used would likely be very different from the device needed by those who are more independent.

Third, there is the low vision traveller who may or may not be using the cane or an optical device. Certainly members of this population who are dependent on the use of the cane and/or an optical device could directly benefit from a new mobility device. More data is needed on low vision travellers in order to determine how many of this population use existing mobility devices and would benefit from a new device.

Fourth, there are the "hidden" blind people. Some are hidden because they need a mobility device but have not been identified for service provision. Others refuse to seek out care from the human services system because of the stigma associated with blindness or because what they really want is to have their sight restored. Although there are no data on these people, it is conceivable that they are significant in number. Some would accept a new type of mobility device if they were informed about it and others would consider a new device if it was less objectionable to them than existing devices.

Age and income are two demographic characteristics of the blind population that need to be considered as well. Prevalence rates of visual impairment increase with age and income is inversely related to prevalence of visual impairment.[3] This would suggest that any new mobility device should meet the needs of elderly people and should be inexpensive and, preferably, third party reimbursable.

Also, O&M specialists are finding that about 50% of the blind people they serve have additional impairments.[6] This is not surprising, considering the prevalence of other impairments among the visually impaired in general.[7] The design of any new device will need to take into consideration the implications of a variety of additional impairments.

Many of the conventional assumptions about independent travel do not fit the situation of the aged blind population, the multiply handicapped blind population, and the institutionalized blind population. Mobility objectives for these populations often revolve around safe and secure ambulation: if the need for a cane is indicated, it is often a short support cane. Whereas independent mobility connotes the ability to travel far and wide, for these people it often entails mastering a select few routes.

Re-Conceptualizing the Mobility Problem

While it is certainly desirable to solve the problem of mobility for blind persons in one "fell swoop", it will likely prove impossible to do for many years. It is worth noting, however, that progress is being made on two research fronts that could have a profound impact on how the mobility problem is solved.

The Totally Blind Traveller

The prospect of developing a prosthetic visual system has always seemed beyond the realm of possibility and practicality but today it seems less so. Over 20 years ago Brindley[8] experimented with implanting a matrix of electrodes in the visual cortex to create a crude sense of vision from phosphenes. Further research on artificial vision was conducted in the 1970's[9]. The National Institute of Health on Neurological Disorders and Stroke has recently developed electrodes that are smaller and require less current and thus can be placed closer together, affording improved resolution.[10] Pursuing a different avenue of research, D. N. Spinelli has demonstrated that the visual cortex can be made directly sensitive to light through the use of non-toxic photo-active chemicals,[11] suggesting the possibility of a non-invasive means of directly stimulating the visual cortex.

The Low Vision Traveller

The Wilmer Eye Institute of Johns Hopkins University and NASA (The John C. Stennis Space Center and other NASA field centers) have recently announced a "Low Vision Research and Development Program" to apply space technology to the problem of low vision:

> "Specifically, NASA is working with the Wilmer Eye Institute to adapt their "robotic telepresence" technology, destined for use in space station

projects, to a portable, head-mounted, computer-controlled image processing and display system." (Joint announcement by Wilmer and NASA)

They will attempt to develop an experimental low vision device or Low Vision Enhancement System that will have the capacity to process a spectacle mounted television image in numerous ways including through magnification, contrast, filtering, and edge enhancement techniques. They will also attempt to remap the visual information from dysfunctional to functional areas of the retina.

These solutions will require years of research and development. For the short-run, it may be more productive to "chip-away" at the mobility problem by isolating and defining very specific subset problems and attempting to solve each subset individually. Whenever possible, solutions to specific problems should be sought that have wide impact, going beyond the blind population, so that commercialization is feasible and at a low cost to the consumer.

Examples of three very different subset problems and solutions follows.

Some Simple Orientation Solutions

For the active and independent blind traveller there may be simple problems that are best solved with simple technical solutions. Returning home from a trip during a snowstorm can present the blind person with the unpleasant situation where usual landmarks are concealed by fresh snow, forcing the traveller to guess at the location of his house. At such times it would be extremely helpful to have a homing device that could lead the traveller to his front door. Many other types of concealed destinations present similar problems that could be solved through the use of a simple device such as an auditory beacon.[12] Prototypes have been developed in the public domain that use components from radio control toys and garage door openers.

The problem of chronic spatial disorientation among institutionalized blind people who have major mental or neurological problems can be conceptualized in a way that lends itself to technical solution:[13]

"When, for example, a severely retarded person who is blind has a chronic disorientation problem, the assistance of the staff in getting to interior destinations in the institution is often a practical necessity, which sets up a cycle of reliance on staff assistance, which fosters dependence, which increases the tendency to rely on staff assistance.

To break the cycle of these residents' total dependence on staff assistance, it would be desirable to have some sort of automated system that could take the place of the staff member in guiding the residents from one interior

destination to another. The system would increase the residents' self-esteem and sense of independence and decrease staff time spent in guiding them. Such a system would also provide instructors with an exciting teaching option. What might emerge, if it were possible to wean the residents from the system, would be a new and potentially powerful teaching tool."

The American Foundation for the Blind has been experimenting with an Auditory Directional System which provides a "musical pathway" for blind persons to follow to get to interior destinations in an institutional setting.[13] The system's components are a compact disc player and a network of speakers, infrared "people-detection" equipment, and a computer controlled speaker-sequencing system.

In mass transit, Electronic Speech Information Equipment that functions as a "talking bus stop" has been developed and tested by the Department of Transport in England in collaboration with the University of Nottingham[14] (in press):

"...Many blind and partially sighted people, who have to rely on public transport, have considerable difficulty in locating the bus stop, establishing whether the bus they need stops at that point, in discovering when the next bus is scheduled to arrive and, finally, in ascertaining which approaching bus is the one that they need.

With these problems in mind the object was to devise a system which would enable a blind or partially sighted person to locate the stop, to activate manually some audible information about services using the stop, scheduled arrival times and the approach of any given bus.

The system was developed around 3 main components. A 'talking' element using voice synthesis, a bus identification unit to 'read' the route number of approaching buses and, linking these two components, a co-ordinating system using a micro-computer.

From the comments made and from personal conversations with users and participating organizations it is clear that the system has generally been very well received and extremely welcome. It is particularly interesting to note that it has demonstrated its usefulness not just as a means of helping blind and partially sighted people but also in giving guidance and reassurance to elderly people and to visitors unfamiliar with the town's bus route and system."

Conclusion

It is often easy for professionals to forget that successful technology development in the field of mobility for blind persons requires the input of many persons with different backgrounds, including blind persons themselves. Blind people need to be consulted to guard against a common

cause of failure: the development of solutions to problems that do not exist. Success also requires that two types of very different professionals cooperate closely: scientists and engineers on the one hand and practitioners and instructors, and especially O&M specialists, on the other hand. The history of technology development in the field of blindness is replete with examples of poor inter-disciplinary cooperation[15] which inhibited the development of very promising technology.

It is difficult to predict how and when the cane, dog, and the array of optical devices that we know of today will be made obsolete. Certainly these mainstays of O&M will be around for many years to come. But if we work together, the day will arrive when truly profound technological solutions to the problem of mobility for blind persons will be available. The will and the desire are there; only the ways and means need to be found.

REFERENCES

1. Blasch, B., Long, R., and Griffin-Shirley, N. 1988 A National Survey of Electronic Travel Aid Users and Former Users. Paper presented at the international conference of the Association for the Education and Rehabilitation of the Blind and Visually Impaired, Montreal, Canada.

2. Uslan, M., Peck, A., and Kirchner, C. 1981 Demand for orientation and mobility specialists in 1980. *Journal of Visual Impairment and Blindness* 75(1):8-12.

3. National Research Council 1986 *Electronic Travel Aids: New Directions for Research*. Washington, D.C.: National Academy Press.

4. Kirchner, C., Stephen, G., & Chandu, F. 1987 Statistics on blindness and visual impairment. Pp. 114-139 in *Yearbook of the Association for the Education and Rehabilitation of the Blind and Visually Impaired*. Alexandria: Association for the Education and Rehabilitation of the Blind and Visually Impaired.

5. National Society To Prevent Blindness 1980 *Vision Problems in the U.S.*. New York: National Society to Prevent Blindness.

6. Uslan, M., Hill, E., and Peck, A. 1989 *The Profession of Orientation and Mobility in the 1980's: The AFB Competency Study*. Unpublished manuscript. New York: American Foundation for the Blind.

7. Kirchner, C. 1985 *Data on Blindness & Visual Impairment in the U.S.*. New York: American Foundation for the Blind.

8. Brindley, G. S. and Lewin, W. D. 1968 The sensations produced by electrical stimulation of the visual cortex. *Journal of Physiology* 196: 479-493.

9. Dobelle, W. H., Mladejovsky, M. G. and Girvin, J. P. 1974 Artificial vision for the blind: Electrical stimulation of visual cortex offers hope for a functional prosthesis. *Science* 183: 440-444.

10. National Institute of Health on Neurological Disorders and Stroke 1988 Stimulating electrodes based on thin-film technology. Fifth quarterly report of contract NO1-NS-6-2309 submitted by the Department of Electrical Engineering and Computer Science of the University of Michigan to the Neural Prothesis Program. Bethesda, Maryland: National Institute of Health on Neurological Disorders and Stroke.

11. Spinelli, D. N. 1988 Is Biovision a Real Possibility by the Year 2020? Paper presented at the conference, Vision Loss: Everybody's Business, Los Angeles, California.

12. Hill, M. 1988 Uses for a simple homing device. *Long Cane News* 7(1): 2-3.

13. Uslan, M., Russell, L., and Weiner, C. 1988 A musical pathway for spatially disoriented blind residents of a skilled nursing facility. *Journal of Vision Impairment & Blindness* 82(1):21-24.

14. Frye, A. 1989 Chapter III in M. Uslan and A. Stern, eds., *The Visually Impaired Traveler in Mass Transit: A Resource Guide*. Unpublished manuscript. New York: American Foundation for the Blind.

15. Spungin, S. 1985 Corridors of insensitivity: Technology and blind persons. *Journal of Visual Impairment and Blindness* 79(3):113-116.

GRADUATES DEGREES IN PHYSIOLOGICAL OPTICS AWARDED BY INDIANA UNIVERSITY

Ph.D. Degrees Awarded
1962 to 1988

Name	*Dissertation Title and Year*
Robert Bert Mandell	Morphometry of the Human Cornea, 1962
John Haas Carter, Jr.	A Servoanalysis of the Human Accommodative Mechanism, 1962
Donald Graves Pitts	Ocular Accommodation in Cat from Electrical Stimulation of the Brain, 1964
William Russell Baldwin	Some Relationships between Ocular, Anthropometric, and Refractive Variables in Myopia, 1964
William Montgomery Lyle	The Inheritance of Corneal Astigmatism, 1965
Brian Ward	Intra-ocular Pressure and the Function of the Neural Elements of the Cat Eye, 1966
Rogers Webster Reading	Extrahoropteral Distance Perception in Tracking Performance, 1968
Arnulf Remole	The Effect of Wave Length on Subjective Intermittence Patterns, 1969
Merrill Emerson Woodruff	Ocular Refractive Trends in the Human Eye Up to Six Years of Age, 1969
Anthony John Adams	Chromatic, Spatial and Temporal Influences on Single Ganglion Cell Responses in the *in vivo* Goldfish Retina, 1970

Paul William Lappin	Ocular Effects from the Helium-Neon CW Gas Laser, 1970
George Chi Shing Woo	The Basis for Panum's Area, 1970
Ben Victor Graham	Color Vision in the Peripheral Visual Field, 1972
Richard Charles Van Sluyters	Binocular Vision in the Rabbit: Cortical Electrophysiology and Behavior in Normal and Monocularly Deprived Animals, 1972
Norman Edward Wallis	Intrastromal Light Scatter at Various Levels of Corneal Hydration, 1972
Roger William Wiley	Studies of the Subcortical and Transcallosal Projections to the Binocular Cortex of the Rabbit, 1973
Morton King Ohlbaum	The Mechanical Resonant Frequency of the Human Eye in vivo, 1973
Perry Speros	Spectral Characteristics of the Early Receptor Potential of the Eye, 1973
Joel Benjamin Spiegler	Photoreceptor Spatial Properties, 1973
James Earl Bailey	Directional Sensitivity of Retinal Receptors, 1973
Joseph Raymond Zahn	Maximum Velocity of the Human Saccade in the Dark, 1975
Richard Dale Hazlett	Physiological Changes of the Cornea Associated with Contact Lens Wear, 1975
Thomas David Williams	Effect of Meridional Disparity on Depth Perception, 1975
Robert William Massof	Wavelength Dependence of Foveal Threshold Characteristics, 1975
Freddy Wilfred Lenox Chang	Determination of the Minimum Threshold of Stereopsis with Random Dot Patterns, 1976
Jerald Wayne Strickland	Electrophysiological Correlates of Binocular Visual-Sensory and Visual-Motor Stress, 1976
David Barry Henson	Investigation into Corrective Saccadic Eye Movements, 1976

George Rexford Courtney	Visual Characteristics of the Institutionalized Mentally Retarded, 1977
Peter Avery Davison	Absolute Saturation Limen as a Function of Luminance and Angular Field Extent, 1978
Gary Lee Trick	The Effect of Wavelength on Binocular Summation, 1978
Kenneth Edward Brookman	Ocular Accommodation in Human Infants, 1980
David Arthur Goss	Effect of Overcorrection of Myopia on Its Rate of Increase in Youth, 1980
James Anthony Worthey	Global Features of Color Rendering and Vision Difficulties Under Fluorescent Lights, 1981
Theanchai Tanlamai	The Comparative Effects of Real, Optical, and Perceived Distance on the Angle of Anomaly in Anomalous Retinal Correspondence, 1981
Richard Lee Martin, Jr.	Photoreceptor Membrane Degradation in the Crayfish: I. Uptake of Ultrastructural Tracers, II. Regulation, 1983
Thomas Redden Colladay	Glutamate Dehydrogenase and Retinal Distrophy in the Peth Rat, 1983
Isaac K. O. K. Kragha	Factors in the Determination of the Age of Onset of Presbyopia, 1985
Jeffery Kirk Hovis	Dichoptic Opponent Hue Cancellation, 1986
E. Peter Osuobeni	Monocular Vernier Acuity in Normally Binocular Monocular and Amblyopic-Mated Eyes, 1986
Michael K. Smolek	Analysis of Ocular Distension via Real-Time Holographic Interferometry, 1986
Mark A. Criswell	Cellular Mechanisms of Movement Dectection and Directionality in the Turtle Retina, 1987
John P. Moxley	Post-Receptor Chromatic Adaptation, 1987
Daphne L. McCulloch	Visual Evoked Potential Measures of Binocularity, 1988

M.S. Degrees Awarded
1956-1988

Name	*Thesis Title and Year*
William Russell Baldwin	A Modified X-Ray Method for the Measurement of the Axial Length of the Living Eye, 1956
Lester Ray Loper	The Relationship Between Angle Lambda and the Residual Astigmatism of the Eye, 1956
Albert Vernon Alder	The Binocularly Induced Phi Movement as a Method of Measuring Binocular Fixation Disparity, 1957
Floyd Marvin Morris	The Influence of Kinesthesis Upon Near Heterophoria Measurements, 1958
John Haas Carter, Jr.	A Review of the Characteristics of Impression Tonometers and Evaluation of the Wolfe Scleral Tonometer, 1959
Donald Graves Pitts	Transmission of the Visible Spectrum Through the Ocular Media of the Bovine Eye, 1959
Ronald Ward Everson	Visual Acuity and Refraction in Relation to Eye Position, 1959
John Reuben Levene	An Evaluation of the Hand Keratoscope as a Diagnostic Instrument for Corneal Astigmatism, 1961
William Montgomery Lyle	Analysis of Monocular Dark Adaptation Measurements Obtained with Alterocular Fixation, 1962
Edward Robert Seefelt	Effects of Initial Spectacle Wearing on Subsequent High School Scholastic Grade Scores, 1962
Marvin Lunsky	Accommodation and Convergence Variations in Low Astigmatism, 1962
Benjamin Kislin	Crystalline Lens Elasticity in Bovine, 1962
Indra Mohindra	The Relationship Between Axial Length of the Eye and Certain Anthropometric Data, 1962

Joel Benjamin Spiegler — Effect of Target Pursuit on Pulfrich Stereophenomenon, 1967

Frederick Van Nus — Ophthalmoscopic and Perimetric Study of the Pecten Oculi in the Avian Eye, 1967

Vigo Hemmershoj Nielsen — Recogition of Elementary Forms in Relation to Visual Acuity, 1968

Morton King Ohlbaum — The Effects of Altitude on Certain Aspects of Visual Performance, 1968

Ralph Swartz — A Comparison of the Ocular Pursuit Movements of Normal and Dyslexic Subjects, 1968

Bertram D. Targove — The Effect of Pilocarpine on the ACA Ratio, 1968

George Chi Shing Woo — Incidence of Amblyopia in Grade School Children in Relation to Age, Sex, and Refractive Error, 1968

Richard Donovan Septon — Apparent Brightness Increase with Yellow Light, 1968

Roger Christian Fitch — Procedural Effects on the Manifest Human Amplitude of Accommodation, 1969

Richard Dale Hazlett — Changes in Corneal Thickness, Corneal Curvature and Corneal Transparency Associated with Contact Lens Wear, 1969

Irving Lawrence Dunsky — An Analsysis of Some Refractive Error Trends in U.S. Air Force Pilots and Navigators, 1969

Wayne Frederick Provines — The Effects of Aging on the Amplitude of Convergence, 1969

Lolita Baluyut Ty — Reliability Analysis of Cheiroscopic Tracing, 1969

Jacob Gershon Sivak — Studies of the Vertebrate Median Eye, 1969

Thomas David Williams — Vertical Disparity in Depth Perception, 1969

Steven Barry Greenspan — Effects of Children's Nearpoint Lenses: Changes in Body Posture and Visually-Centered Performance, 1970

Louis Vito Genco — The Effect of Altitude on Critical Fusion Frequency, 1970

Edwina Alice Challinor	The Relative Role of Physical Features of Spectacles as Factors in Wearing Comfort, 1970
Charles Junior Archibald	The Pulfrich Stereophenomenon Measured Lateral to the Point of Fixation, 1971
David P. Austen	The Gross Anatomy of the Human Ciliary Muscle, 1971
Rosa Isabel Revuelta	Effect of Caffeine and Caffeine Combined with Vitamin A on Dark Adaption, 1971
George William Mikesell, Jr.	Alterations in Eye Movement and in the Lateral Phoria as a Result of Hypocapnia Produced by Voluntary Hyperventilation, 1971
Daniel Robert Gerstman	Variation of Corneal Thickness in Clean and Polluted Air, 1971
Ian Laurence Bailey	Monochromatic Aberration of the Eye With Contact Lenses, 1971
James Thomas Gallagher	The Effect of Gravity Direction Change on Amplitude of Accommodation, 1972
John Frederick Amos	Early Receptor Potential Amplitude as a Function of Intensity in Three Species of Monkeys, 1972
David Dudley Glick	The Effect of Disability Glare on the Visibility of Line Gratings, 1973
Charles Leslie Haine	Laser Meridional Refractometry, 1975
Kenneth Edward Brookman	Comparative Morphology of the Vertebrate Cornea: A Light and Phase-Contrast Microscopic Study, 1975
Eugene Jon Potvoricky	Interpupillary Distance Measurement, 1975
Stanley D. Miller	The Efficacy of Emergency-Use Highway Warning Signals, 1975
Raymond Alan Applegate	Contrast Sensitivity and the Refractive State, 1976
Donald Thomas Lowman	Refractive Errors in the Developmentally Disabled Academic Underachiever, 1979

Mark M. Uslan

Cane Technique: Toward More Complete Path Coverage, 1977

Theanchai Tanlamai

Prevalence of Monocular Amblyopia Among Anisometropes, 1978

Stella Tatonye Briggs

Efficacy of a Thermal Disinfecting Unit for Soft Contact Lenses, 1979

Philip Courtney De Santis

Aldehyde Fixation for Corneal Adenosine Triphosphatase, 1979

Edwin Cochran Marshall

A Demographic and Epidemiological Study of Two Clinical populations, 1979

Robert E. Miller, II

The XM-30 Military Protective Gas Mask: Its Effect on Ocular Environment and Corneal Thickness, 1982

Carol A. Westall

Fixational Eye Movements and Autokinesis in Amblyopia, 1982

David J. Walsh

Peripheral Visual Acuity: Oblique and Meridional Effects, 1985

Richard J. Dennis

The Effects of Long-Term Monocular Occlusion on Vernier Threshhold: Elasticity in the Young Adult Visual System, 1986

	DATE DUE		